SENSITIVITY OF HIGH-RESOLUTION WRF FORECASTS TO A SINGLE RADIOSONDE IN A DATA SPARSE REGION

A Thesis

by

MARK DAVID BENOIT

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Chris Nowotarski
Co-Chairs of Committee,	Don Conlee
Committee Members,	Istvan Szunyogh
	Huilin Gao
Head of Department,	Ping Yang

December 2016

Major Subject: Atmospheric Sciences

ABSTRACT

Radiosonde observations (RAOBs) are relatively sparse over central Texas. The closest RAOB launch site to College Station, Texas is in Fort Worth, Texas, about 250 km away. On-demand soundings were launched by Texas A&M University students in high-impact situations. Both the local NWS offices and the SPC requested RAOBs. These observations had value to forecasters in convective and winter weather situations.

The purpose of this research was to find the value of on-demand RAOBs on a high-resolution NWP forecast for College Station. DA was done with 29 RAOBs using WRF, but also incorporated other observations from MADIS. In total, there were 116 simulations since four simulations were done for each RAOB. Using weather model analyses and observations from the EOL, MADIS, and NOAA, the simulations were evaluated. DTC-MET and SPoRT-MET tools utilized the datasets to provide verification.

In some cases, DA of a single RAOB produced modest, positive impacts on the WRF forecast. Precipitation characterization and high precipitation amounts were improved in convective cases, while surface and low-level temperature forecast improvements were seen in short-range forecasts for winter cases. Benefit was spatially limited to areas near College Station, and was further limited when MADIS observations were assimilated. Future work supports RAOB launches in high-impact situations; however, real-time DA of these RAOBs is not a high priority.

ACKNOWLEDGEMENTS

My advisor Dr. Chris Nowotarski deserves the biggest acknowledgement. Thank you so much for guiding me through the project and giving me what I need to accomplish the project. Dr. Don Conlee, thank you for being my teacher in undergrad and helping me do much more. My two other committee members, Dr. Istvan Szunyogh and Dr. Huilin Gao, were integral parts of the process as well. Thank you for advice and setting aside time from your busy schedules.

Others also deserve mention for supplemental help along the way. Dr. Jon Poterjoy at NCAR provided some helpful MatLab scripts for RAOB DA, Lance Wood provided references to forecast verification scripts, and Christine Paschal analyzed the RAOBs before they were used in WRF-DA. Two help desks were invaluable to troubleshooting, which were the WRF help desk at NCAR, and the DTC-MET help desk. Support was also received from so many colleagues, friends, and family. For these people, I am grateful.

NOMENCLATURE

DA	Data assimilation
CSI	Critical Success Index
FAR	False Alarm Rate
MADIS	Meteorological Assimilation Data Ingest System
NWS	National Weather Service
NWP	Numerical Weather Prediction
POD	Probability of Detection
RAOB	Radiosonde Observation
RAP	Rapid Refresh
RH	Relative Humidity
RMSE	Root Mean Squared Error
RTMA	Real-Time Mesoscale Analysis
SPC	Storm Prediction Center
SOUP	Student Operational Upper-air Program
QPF	Quantitative Precipitation Forecast
WRF	Weather Research and Forecasting

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION

1.1 Central TX Upper-level Observation Hole

Central TX has sparse observations at upper-levels of the atmosphere as noted by the lack of radiosonde observations (RAOBs) over the area. The closest launch site is in Fort Worth, TX, which is about 250 km from College Station. Figure 1.1 shows the distance between balloon sites, where CLL is the location of Texas A&M University. The vertical profile of the atmosphere can be roughly diagnosed by model interpolation or advection from other observed areas, but a RAOB provides the best high-resolution vertical measurement of the atmosphere at a specific location.

The atmospheric profile can be estimated from satellites, but retrievals are not as accurate as in situ observations. They are not sensing actual atmospheric variables (Divakarla et al., 2006), but their retrievals are analyzed by algorithms to derive the atmospheric profile. Infrequent times are also common since many retrievals are taken by polar-orbitting satellites, which only fly over a certain regions and are available once or twice a day. In addition, the vertical resolution of the satellite level observations such as those from the Atmospheric InfraRed Sounder (AIRS), have a vertical resolution of 1-2 km (Goldberg et al., 2003), which is quite low given how much the atmosphere can vary over that depth.

Many of these measurements from satellites are used for real-time forecasting and numerical weather prediction (NWP), and are beneficial. Their limitations, however, should be understood. Wind retrievals from satellites, for example, are much less accurate than RAOB winds. The assigned height of the wind retrievals can have large errors. For instance, (Deb et al., 2008) found that the mean absolute error in vertical height was 27 hectoPascals (hPa) for water vapor imagery. In contrast,

RAOBs measure winds to the nearest several hPa.

RAOBs provide forecasters with valuable observations that can be plotted on thermodynamic diagrams such as the Skew-T Log-P diagram, but they are known to be important data sources for NWP as well. For example, an ensemble reanalysis done without radiosondes showed errors in 500 hPa height that were similar to 48-hour forecast errors (Whitaker et al., 2004), which are on the order of 40 m in the extra-tropics. A single observation at College Station with a RAOB would likely not deviate 40 m from a model's guess of 500 hPa height, but any deviation in height could affect the model forecast of the upper-level atmospheric height in the region of central TX.

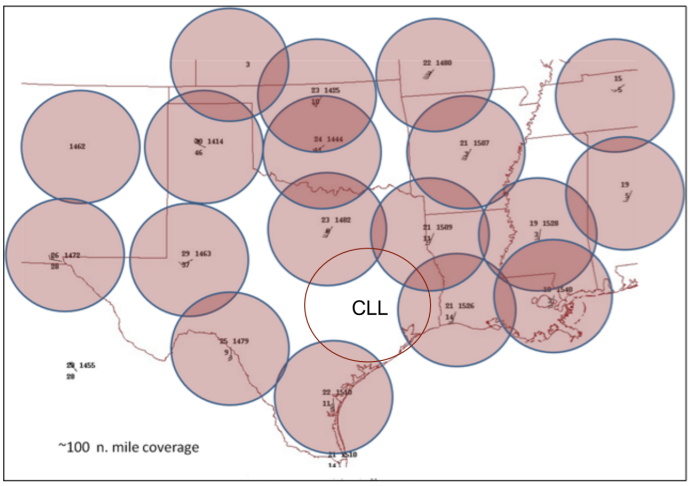

Figure 1.1: Upper-level observation hole over Central TX. CLL is not a standard launch site, but is in an area of sparse RAOBs.

1.2 SOUP Program at Texas A&M and NWS Use

The Student Operational Upper-air Program (SOUP) is a program run by the Department of Atmospheric Sciences at Texas A&M University. They provide on-

demand RAOBs to three surrounding National Weather Service (NWS) offices, and the Storm Prediction Center (SPC). They are launched during active convective or winter weather situations to fill a gap in observations at upper-levels of the atmosphere. When an office contacts Texas A&M about launching a balloon, students are mobilized to launch at the requested time. These soundings have proven useful to forecasters, as they are often cited in the forecast and mesoscale discussions. This RAOB data undoubtedly gives them a better knowledge of forecast parameters, and helps them better predict convective, severe, and winter weather in southeast TX.

Several parameters that help forecasters predict convective and severe weather are Convective Available Potential Energy (CAPE), Convective INHibition (CINH), wind shear, and other computed quantities from a Skew-T Log-P diagram. CAPE diagnoses whether there is enough energy in the environment and affects the updraft strength of storms, which is important for prediction of severe weather, CINH levels are very important to know for the initiation of convection, and wind shear especially can be important in the lower levels of the atmosphere because it affects the storm relative helicity and rotation. For winter weather, the freezing level and the vertical profiles of temperature and moisture are important. Freezing level and moisture in the vertical are essential for determining the originating precipitation type, while the vertical temperature profile is important for diagnosing precipitation type change. Freezing level is also useful in convective situations since it signifies the level at which hail starts melting.

1.3 Review of Observation Impacts and Single RAOB Assimilation

This study hypothesizes that just as these supplemental RAOBs are useful to forecasters in a data sparse upper-air region, there is potential benefit to NWP; particularly in a high-resolution convection allowing model around College Station.

State variables such as temperature (T), pressure (P), water vapor (q_v), and wind are predicted by NWP models. Different values of these variables will not only affect quantities such as CAPE and CINH, but quantities related to winter weather as well. Assuming that a single RAOB in a data sparse region can improve the model analysis (i.e., initial conditions), it is expected that the forecast could also improve.

It is well known that RAOBs, in general, are useful to NWP. A recent study shows that they have a positive impact on short-range forecasts ($\leq$ 24 hours) in the Rapid Update Cycle (RUC) Model, being second only to aircraft observations overall (Benjamin et al., 2010). Figure 1.2 from Benjamin et al. (2010) demonstrates the impact of RAOBs as a whole on relative humidity (RH) across the United States. Results are based on data denial experiments and use RAOBs as truth, where control experiments included all data types. The numbers in the figures are then the root-mean-square error (RMS) of a specific data type minus the Control RMS. The left figures show the total impact of each observation subset on a specific date, whereas the right figures show impact at 3, 6, and 12 hour forecast intervals.

Targeted observations can add some skill to forecasts, but improvements are not always seen. Langland (2005) reviewed different studies and found that impact was variable, as 70% saw improvement. In some cases, short-range forecasts were improved by 50%. The majority of the supplemental data in these studies was targeted dropsonde data. Satellite observations are becoming increasingly more valuable (Bouttier and Kelly, 2001; Kelly et al., 2004), but impacts are still seen for RAOB data as AMSU radiances and radiosondes provide a similar reduction in forecast error (Gelaro et al., 2010). Another study done by Kelly et al. (2007) used limited observations over the Atlantic and Pacific, with impacts local and downstream of those observations. A single RAOB has little impact on a global or regional model forecast, however, this study is focused on its local impact in a

high-resolution, convection allowing model. The RAOB will be assimilated it into a high-resolution model and for a data sparse upper-air region. Since College Station is in a significant upper-level observation hole, it is hypothesized that a single RAOB will add forecast skill to this region.

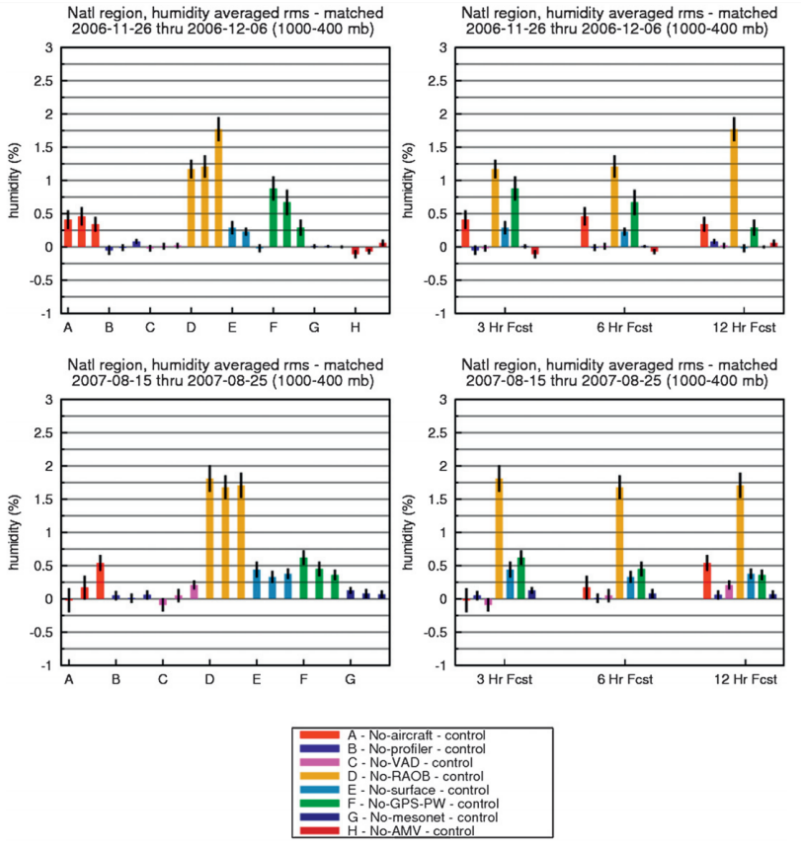

Figure 1.2: Data denial experiments with different instrument types. Left panels show overall impact of each instrument for different dates, while the right panels show impact at 3, 6 and 12 hour forecasts for those dates. RMS error based on RAOBs (Benjamin et al., 2010).

1.4 Data Assimilation

Data Assimilation (DA) combines a short term forecast from a NWP model and observations to obtain a best guess of the atmospheric conditions at the initial start

time of a forecast. Better initial conditions should lead to better forecasts; however, observations and forecasts each have errors associated with them, so estimating this error is important. DA aims to assign the best error estimates to observations and NWP forecasts, and combine them to better estimate the initial state of the atmosphere.

RAOBs compose a significant portion of total data assimilated (Huang and Vedel, 2003). A typical sounding has data from the surface to 100 mb or higher. In this project, RAOBs from Texas A&M University in College Station, TX are assimilated into a well-known NWP model, the Weather and Research Forecasting model (WRF) (Michalakes et al., 2001) using the three-dimensional variational data analysis (3D-Var) approach, to produce forecasts that will be compared to other simulations that have different initial conditions (ICs).

1.5 Project Scope

As stated above, this project aims to determine the impact of RAOBs on high-resolution NWP forecasts in College Station, TX and regions near it. There are many predicted weather variables and derived outputs that affect official weather forecasts, and this affects government and general populous decisions. One such output is simulated radar reflectivity, which gives information not only about precipitation intensity and timing, but also can be used to diagnose convective mode and severe weather threats. Other outputs like temperature are beneficial, especially in winter weather situations; a few degrees can alter precipitation type from rain to freezing rain to snow.

The main research question motivating this work is then, does a single RAOB in College Station have significant NWP value? Experiments will assimilate these RAOBs into a NWP model, along with other assimilated observations from the

Meteorological Assimilation Data Ingest System (MADIS), to examine their benefits. It is hypothesized that some benefit will be seen for the single RAOB assimilation, but when MADIS data is assimilated, the RAOB will have a smaller impact. Verification of the simulations will be done for radar reflectivity and precipitation accumulation with radar data from the Earth Observing Laboratory (EOL), and temperature, dewpoint temperature, relative humidity, and wind is verified with observations from MADIS, the Real-Time Mesoscale Analysis (RTMA), and the Rapid Refresh Model (RAP) at NOAA. These datasets are used to determine local NWP value of an additional RAOB in a data-sparse region. It is hypothesized that most NWP value will be localized around College Station and downstream. Since MADIS data is comprised of a significant amount of data, its assimilation will likely have a larger impact on a forecast than a single RAOB.

2. BACKGROUND AND LITERATURE REVIEW

2.1 NWP and Data Assimilation Methods

NWP is an initial value problem, and as such, relies on an estimate of the atmosphere to get a weather forecast. Richardson (1922) and Charney et al. (1950) were both instrumental to NWP. They performed hand interpolations of available observations to grid points, but this was very time consuming (Kalnay, 2003). It soon became apparent that an objective analysis was needed to fit the data to grids, and methods were developed (Panofsky, 1949; Gilchrist and Cressman, 1954; Barnes, 1964). Outlined in the following paragraph is a brief history of the NWP and different types of DA.

Various parts of NWP improved with objective analyses (Cressman, 1959), but it was clear that a first guess of the atmosphere at all grid points was necessary to generate initial conditions for a forecast (Bergthórsson and Döös, 1955). Poor observation coverage made it necessary to develop the background or first guess estimate of the atmosphere. Previously, an analysis of the atmosphere was determined from climatology, but now a combination of climatology and first guess could be used. This led to empirical methods of DA such as the successive correction method (SCM), which blended developed climatology and the first guess (Bergthórsson and Döös, 1955; Cressman, 1959), and nudging, which changes the initial solution towards observations (Hoke and Anthes, 1976; Kistler, 1974). Other forms of DA combine observations and a short-term NWP forecast statistically, which is common today. Two common statistical methods are known as 3-dimensional and 4-dimensional variational (3-D Var and 4-D Var) DA. Adaptive methods are also widely used, such as the ensemble Kalman filter (EnKF) and the extended Kalman filter (EKF)

techniques (Huang and Vedel, 2003), but are computationally expensive. Since 3-D Var is widely used in NWP and fairly accessible, it is utilized for this research.

In some respects, 3-D Var and 4-D Var are similar DA methods. They both operate with similar equations over a three-dimensional domain, but 4-D Var can assimilate data in time as opposed to 3-D Var, which operates at a single time. Thus, if an observation was not taken at the analysis time, say at 1805Z, 4-D Var would assign a different weight to it. 3-D Var assumes that the observation happened at 1800Z regardless of its time stamp. As expected, the systems have slightly different equations, which are referred to as cost functions. It describes how much possible error there is in an analysis, so just as the business world wants to minimize cost, possible error of the atmospheric analysis should be at a minimum. 4-D Var will have a lower cost function because it weights observations with time, but it is also much more computationally expensive. For reasons of computational expense and the fact that the sensitivity to a single observation is being tested, 3-D Var will be used in these experiments. The cost function for 3-D Var, taken from Huang and Vedel (2003), is

$$J = \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x_b})^T \boldsymbol{B}^{-1}(\boldsymbol{x} - \boldsymbol{x_b}) + \frac{1}{2}(H(\boldsymbol{x}) - \boldsymbol{y})^T \boldsymbol{R}^{-1}(H(\boldsymbol{x}) - \boldsymbol{y}) \tag{2.1}$$

where $\boldsymbol{x}$ is the state vector, the analysis. $\boldsymbol{x_b}$ is the first guess of the atmosphere from a short term NWP model forecast, $\boldsymbol{B}$ is the error covariance matrix of the background field, H is an observation operator, which maps model space to observation space, $\boldsymbol{y}$ is a vector containing all observations, and $\boldsymbol{R}$ is the error covariance matrix for observations. Minimizing this function, done by taking its derivative, will give the most likely state of the atmosphere, which is referred to as the maximum likelihood approach by Kalnay (2003).

Several variables in this equation need to be further explained, which are H, $\boldsymbol{B}$, and $\boldsymbol{R}$. H performs an interpolation on the model field to the location of the observation, as well as any other transformations such as converting certain quantities like water vapor to dew point. $\boldsymbol{B}$ has non-diagonal elements related to the background field, which are important to the data analysis. These elements will determine how much the observations will be spread spatially around the observation, both vertically and horizontally (Huang and Vedel, 2003). Not only this, but modification of variables such as temperature will negate changes in pressure, wind, etc. This is one of the most important elements to a data assimilation system. Though it has the most impact on results, it is one of the hardest things to estimate (Navon, 2009; Kalnay, 2003).

Three methods to estimate $\boldsymbol{B}$ are outlined by (Huang and Vedel, 2003), the Hollingsworth Lönnberg method (Hollingsworth and Lönnberg, 1986; Lönnberg and Hollingsworth, 1986), the National Meteorological Center (NMC) method (Parrish and Derber, 1992), which is the most popular and universally adopted method, and the analysis ensemble method (Fisher, 2003a,b). The first method (Hollingsworth and Lönnberg, 1986; Lönnberg and Hollingsworth, 1986) is able to determine the background error and observation error correlations. This method compares forecasts to actual observations to obtain error matrices. These error matrices are based on observation type. It is not always practically possible since it assumes a homogeneous observation network. The second method assumes that background error covariance is proportional to correlations of forecast differences between forecasts of different age, but valid at the same time. Often, forecasts of 24 and 48 hours are used, but it is possible to use other times. The last method uses offsets between short-term forecasts made from different perturbed analyses to estimate $\boldsymbol{B}$. This static covariance matrix

represents the forecast errors, but it may underestimate the data impact because it has deficiencies in flow dependency (Buehner, 2005).

$\boldsymbol{R}$ is a diagonal matrix, meaning that observation errors are uncorrelated. This makes sense since observations do not depend on what is measured somewhere else. $\boldsymbol{R}$ tries to represent the errors of observations, and can roughly be estimated by the statistics of the offsets between the observations and NWP model analyses. Observations errors can be from satellite retrievals, standard instrument error, poor instrument placement, or other things like infrequent instrument calibration.

Once there is an estimate of $\boldsymbol{B}$ and $\boldsymbol{R}$, the cost function (2.1) can be minimized. This is a multi-step problem. First, $\boldsymbol{x}$ is considered equal to $\boldsymbol{x_b}$. The gradient of that cost function is found by taking its derivative, and then $\boldsymbol{x}$ is changed such that 2.1 is a lower value. This step is repeated until 2.1 is at a minimum, or a number of iterations can be specified. The value of the state vector $\boldsymbol{x}$ that minimizes the cost function is used as the model initial condition, the analysis. The analysis will suffer if $\boldsymbol{B}$ or $\boldsymbol{R}$ are poor estimates.

Ensemble and extended Kalman filter techniques are adaptive data assimilation techniques that are computationally expensive, but they offer some advantages. The EnKF has error covariances that are based on an ensemble of model states (Evensen, 2009). $\boldsymbol{B}$ has to be calculated for each model run from many members of an ensemble, taking much more time than a typical DA cycle. The data is also randomly perturbed for the individual model runs (Kalnay, 2003). The EKF predicts error covariance with a linearized and approximate equation (Evensen, 2003) and $\boldsymbol{B}$ is advanced using the model itself (Kalnay, 2003). Studies show that ensemble and and extended Kalman filter methods perform similarly (Madsen and Canizares, 1999).

Each DA method has strengths and weaknesses. 3-D Var is computationally cheap compared to adaptive methods and widely used, but the assumption that $\boldsymbol{B}$ is

constant can be a problem if the most accurate forecast is wanted. This is a poor assumption because a model has different deficiencies based on weather type and season (Kalnay, 2003). Adaptive methods, on the other hand, have big advantages with respect to $\boldsymbol{B}$ in 3D-Var as it changes when the model steps forward in time. As stated above, wind has a flow-dependent covariance matrix, and this is accounted for in the adaptive methods. 3-D Var data assimilation is used for these WRF experiments for simplicity and lower computation cost. Even though the other methods perform better, this study simply utilizes 3-D Var to test the sensitivity of observations to a model forecast. In addition, this method is operational at meteorological centers around the world.

2.2 High-resolution NWP

Lewis Fry Richardson was the pioneer of NWP with his book ‘‘Weather Prediction by Numerical Process’’ in 1922 (Richardson, 1922). His calculations were impractical until later in the century, but the methodology he proposed is largely used today (Lynch, 2006). Computing power has increased significantly since data assimilation began in the mid-1900s, allowing high-resolution NWP and DA. The horizontal grid spacing in the National Meteorological Centers barotropic model was 381 km in the 1950s (Mass et al., 2002), and today, the Global Forecast System (GFS) has a horizontal resolution of around 13 km for the whole globe. Smaller regions, like the contiguous United States, have models such as the National Oceanic and Atmospheric Administration's (NOAA) High-Resolution Rapid Refresh (HRRR) model, which has a horizontal resolution of 3 km.

It is much more expensive to run high-resolution simulations, which begs the question, is higher horizontal resolution better? ICs and boundary conditions (BCs) for high-resolution models are taken from global and regional models, which does not

add information to the model unless you assimilate observations on a smaller scale. Mass et al. (2002) shows that a grid spacing of 12 km might be an ideal resolution as verification scores from a 4 km model with the same input are indistinguishable. The study did, however, find that mesoscale features looked more realistic in the 4 km simulation than in the 12 km simulation. A more recent study evaluated the horizontal grid spacing for the WRF for convection-allowing forecasts at 1 km & 4 km, and no significant benefit was achieved by increased resolution (Johnson et al., 2013).

Nevertheless, since computing has made high-resolution modeling possible, many high-resolution models are operationally run. This has the advantage of resolving convection, which takes place on scales of several kilometers, or even less. Therefore, parameterization of convection, which is the simplification of small scale processes, is not done because the scales are small enough in the model to directly compute the quantities needed. The density of observations has increased in many countries, including the U.S.A., so high-resolution data assimilation is also possible.

Some observation types like GPS precipitable water have become more prevalent in the last couple decades and are used in the operational Rapid Refresh (RAP) model. Assimilation of GPS precipitable water in the Rapid Update Cycle (RUC) model showed modest (5-7%) forecast improvement in 3-hour RH forecasts below 500 hPa (Gutman et al., 2004). This is helpful for real-time forecasts because more moisture in the lower atmosphere affects rain amounts, which can affect flood forecasts. As seen from Figure 1.2 earlier, RAOBs also impact RH prediction. Many studies that do high-resolution model runs and assimilation of observations look specifically at how precipitation in convective storms is affected. One such study assimilated hypothetical high-resolution surface observations with an EKF system. Assimilation of all standard surface variables (wind, temperature, etc), had a significant and

positive impact on forecasted convection (Dong et al., 2011). This was done with an EKF, so results are not directly applicable to other DA methods, but it shows that DA is valuable to high-resolution forecasts.

Another data type used by the DA community is from remote sensing. Radar in particular is often used for convective weather modeling. Short-term forecasts improve with radar and conventional DA, but long-term forecasts with radar DA are not always better, even considering subjective verification of reflectivity (Kain et al., 2010). Hou et al. (2013) did model runs with WRF and ARPS (Advanced Regional Prediction System) 3D-Var packages to evaluate the impact of the DA on heavy rainfall prediction over southern China with a 4 km, convective scale grid resolution. Their control experiment assimilates radar data, but no surface or radiosonde observations. Other model runs incorporate surface observations, radiosonde observations, or both. Overall, surface DA offered a slight positive impact on rainfall location and forecasted surface variables, while radiosonde DA improved the quantitative precipitation forecast (QPF) skill in terms of improving rainfall position accuracy and reducing rainfall over-predicition. Incorporation of all data types had more positive impacts on the forecast than either data type alone.

Similar to the study in China, heavy rainfall events were examined over central Korea using WRF and its 3D-Var system assimilating radar and surface data. Threat scores for rainfall thresholds above 30 mm improved when all data types were assimilated (Ha et al., 2011). Figure 2.1 shows the threat scores for different precipitation thresholds when surface data from automatic weather stations (AWS) and radar data were assimilated. AWS data improved the simulation more than radar by itself. Besides using radar data, a more recent study by Yesubabu et al. (2015) over Saudi Arabia assimilates conventional surface and upper-level observations along with satellite radiances in the WRF 3D-Var system. Assimilation of all observations

contributed to an improved forecast of extreme rainfall events at lead times of 48, 72, and 96 hours.

While these studies show that DA with radar data improves forecasts, 3D-Var DA with radar observations does not always produce better results, as statistics vary in other studies on a case-by-case basis (Sun et al., 2012). If more accurate QPFs are wanted, 4D-Var in general performs well (Sun and Wang, 2013), with higher skill scores than 3D-Var. In summary, 3D-Var has some promise for high-resolution model forecasts due to its lower computation cost, but it also has limitations (Rennie et al., 2011).

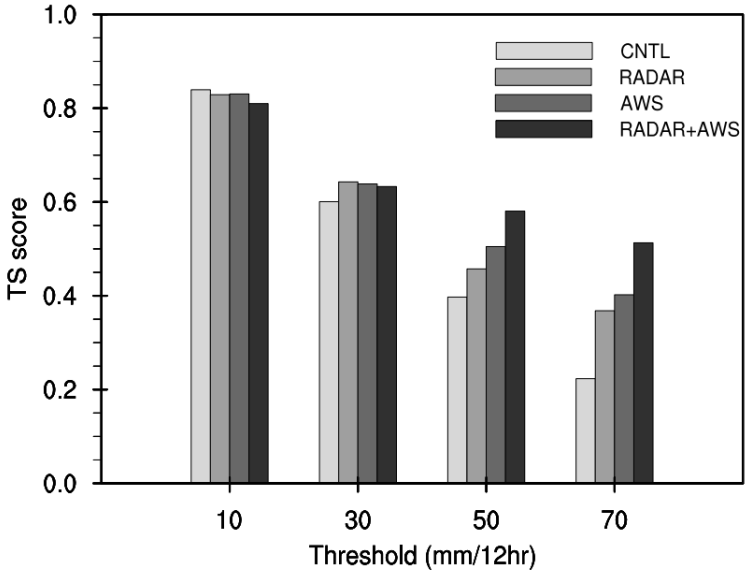

Figure 2.1: Threat scores for different DA schemes (Ha et al., 2011)

3. DATA AND METHODS

3.1 Radiosonde Data

It is necessary to explain the weather balloon instruments used in these model runs as well as details related to the situations for which they were used. The data collected at Texas A&M was from International Met Systems (Inter-Met) iMet-1 radiosondes (http://intermetsystems.com/index.php/products/imet-1). Components of the instrument include a glass bead thermistor for fast response temperature, a thin-film capacitor for relative humidity (RH) measurements, a pressure sensor, and a 12-channel GPS receiver for obtaining height, wind direction, and wind speed. Some radiosondes used in this study had larger capacitors for measuring RH, but the majority had the thin film capacitors. Before the advent of GPS, all radiosondes had pressure sensors, but many do not have a pressure sensor now. Section 3.3 will review the different pressure calculation methods for new sondes lacking a pressure sensor. Of the balloons launched, 23 were launched in advance of a potential convective weather event, and 6 for a potential winter weather event. Below is Table 3.1 that provides a summary of the RAOBs collected at Texas A&M, along with brief descriptions, while Figure 3.1 shows the time frequency of the launches.

Table 3.1: RAOB launch details

Time(YYYYMMDD_HHZ)	Type(Convective or Winter)	Weather Synopsis
20110520_18Z	Convective	500 mb low over the Rocky Mountains. Strong low-level jet (LLJ) over TX.
20120415_15Z	Convective	500 mb trough to the west, with a strong low-level jet.
20121126_23Z	Convective	Shortwave at 500mb stretching from NM into Mexico.
20130210_06Z	Convective	500 mb trough approaching from the west along the NM and AZ border.
20130320_00Z	Convective	Weak trough at 500 mb extending from the Great lakes and down into TX.
20130410_12Z	Convective	Deep trough at 500 mb extending from the Rockies to western Mexico.
20130509_18Z	Convective	Weak shortwave at 500 mb in Mexico. Moderate LLJ of 20-30 knots in TX.
20131126_09Z	Winter	500 mb upper-level low over the north TX panhandle, moving to DRT.
20131205_20Z	Winter	Upper-level low at 500 mb to the north, with shortwaves over Mexico
20131208_00Z	Winter	500 mb trough over the western US, with shortwaves in Mexico.
20140124_00Z	Winter	Several shortwaves at 500 mb in Mexico
20140328_18Z	Convective	500 mb trough from the Rockies to ELP and Mexico.
20140403_21Z	Convective	Trough at 500 mb along the AZ and NM border. LLJ of 55 knots over TX.
20140427_18Z	Convective	Upper-level low at 500mb and trough through NM and TX. LLJ of 55 knots.
20140729_15Z	Convective	500 mb ridge in place, but WAA at 850 mb notable.
20141121_17Z	Convective	500 mb trough in southern CA, but 850 mb WAA is substantial.
20141121_21Z	Convective	Same as the previous case.
20141122_16Z	Convective	Trough at 500 mb moves over central Mexico, with a LLJ of 40 knots.
20150224_00Z	Winter	Small sounding to $\sim$650 mb. Some shortwaves at 500 mb in Mexico.
20150305_06Z	Winter	Deep trough at 500 mb from the Great Lakes and extending into NM
20150417_18Z	Convective	Upper-level low at 500 mb over the four-corners region.
20150424_20Z	Convective	500 mb trough from AZ into central Mexico. LLJ of 50 knots in TX.
20150427_00Z	Convective	Low at 500 mb over NM and a trough through TX and Mexico.
20150505_21Z	Convective	Trough at 500 mb along the AZ and NM border and into eastern Mexico.
20150701_20Z	Convective	Ridge at 500 mb, but a weak upper-level low in Mexico.
20160330_18Z	Convective	Upper-level low at 500 mb over the four-corners. LLJ of 50 knots.
20160418_00Z	Convective	500 mb low east of the four-corners. Significant flooding in Houston.
20160427_02Z	Convective	Trough at 500 mb and stretching from WY into west TX.
20160429_20Z	Convective	Low at 500 mb from southern CO to Mexico. LLJ of 40 knots over TX.

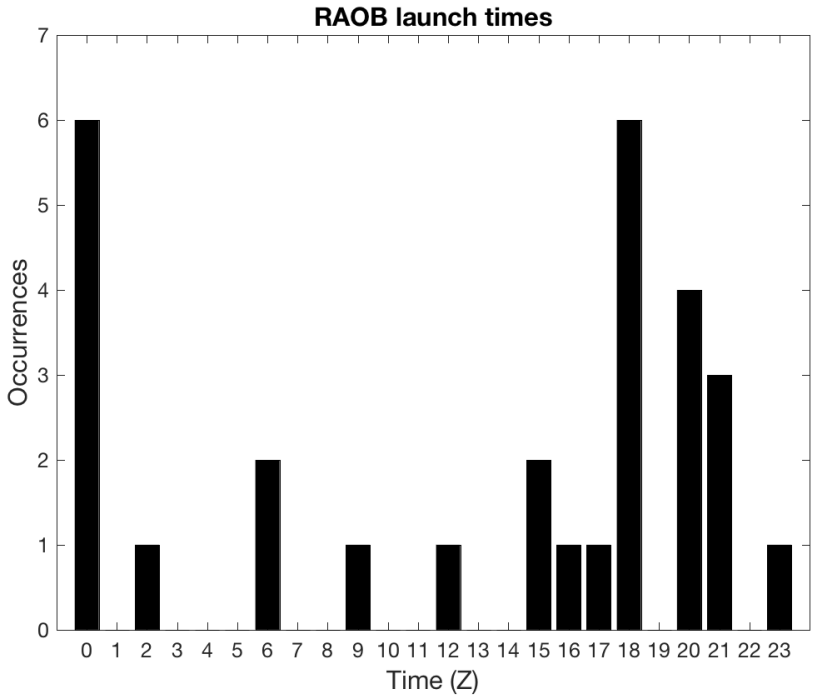

Figure 3.1: RAOB launch time frequency.

3.2 Other Assimilated Data

Radiosonde data assimilation is expected to impact local weather forecasts for College Station, but other widely available data will also be assimilated to emulate a real-time DA. This data is expected to have more forecast impact than a single RAOB. The MADIS (Miller et al., 2005a) will be used in this study, and is comprised of weather data from NOAA and non-NOAA providers such as departments of transportation, universities, volunteers, and the private sector. Observations from MADIS that are used include METAR reports from Automated Surface Observation Systems (ASOS) and Automated Weather Observation Systems (AWOS) at airports, which includes Surface Aviation Observations (SAO), aircraft based reports that included vertical profiles from the planes at airports (AirReps), mesonet observations,

which included NOAA and non-NOAA providers, maritime observations, RAOBs, NOAA profiler networks (NPN), multiple agency profilers (MAP), and satellite wind observations from IR and visible satellite at different levels in the atmosphere, known as high-definition winds (HDW).

3.3 Data Quality and Limitations

RAOB data at Texas A&M were collected using Inter-Met radiosondes and their receiving systems. Geopotential height measurements were obtained in two ways from the radiosondes. The first way utilizes the pressure sensor on the radiosonde along with temperature and humidity to get heights according to the hypsometric equation, which is below.

$$h = z_2 - z_1 = \frac{R_d \overline{T_v}}{g} \ln\left(\frac{p_1}{p_2}\right) \tag{3.1}$$

z_2 and z_1 are the geopotential heights of two pressure surfaces, R_d is the dry air gas constant, $\overline{T_v}$ is the average virtual temperature of that layer, g is gravity, and p_2 and p_1 are the pressure at the respective z heights. The second method still uses the hypsometric equation, but instead of calculating height, it uses GPS heights, temperature, and humidity to obtain pressure. The second method is widespread now, as GPS is much more accurate now. As a result, manufacturers have been producing radiosondes without pressure sensors. Both methods have downfalls and errors, the first of which will be discussed now.

A previous radiosonde model called RS-80 is made by Vaisala and has a pressure sensor much like iMet-1. At altitudes of 26 km, evaluations of iMet-1 and RS-80 radiosondes show similar variable pressure offsets of -0.65 hPa and -0.55 hPa respectively (Stauffer et al., 2013). If a pressure uncertainty of 0.4 hPa +/- 0.2 hPa is assumed above 20 km for a radiosonde, geopotential height differences can be 42

+/- 24 m (Inai et al., 2015) and increase with height. Previously, pressure bias of the RS-80 were found to be -0.5 hPa above 15 km and statistically significant (Inai et al., 2009). These findings gave geopotential height differences of 100 m or more near 20 km. Not only do these errors and bias affect geopotential height along with other meteorological variables like temperature and humidity, but they also affect ozone (O_3) measurements in the stratosphere Stauffer et al. (2013).

The second method does not use a pressure sensor from the radiosonde, but still needs a pressure at the ground. Starting with a good reference barometer at the ground, there are still some uncertainties in calculations. Poor GPS reception, temperature, and humidity measurements will all affect the calculations of pressure that are made with the hypsometric equation. Inai et al. (2009) showed that geopotential heights obtained with the second method are more accurate. Christine Paschal was a Research Experience for Undergraduates (REU) student at Texas A&M in the summer of 2015 and also investigated this, with the same results. Operationally, many weather service offices use radiosondes without pressure sensors, and the RAOBs used in this study will use the second method of calculation.

MADIS data goes through an extensive quality control procedure (Miller et al., 2005b) before it is available for data assimilation. It assigns flags to the data based on their quality. Unfortunately, software that converts this data format for use in the WRF does not translate the flags into the new data format, so there is potential for bad data to be used in the model. Section 3.5 will talk about WRF's DA system and the quality control that is designed to keep poor quality data from assimilation.

3.4 WRF Model

The WRF model with the Advanced Research WRF (ARW) core, version 3.6, is used for this study. It is designed for both research and operational applications

to advance the understanding and prediction of mesoscale weather. The reader is referred to Skamarock (2008) for more details on model formulation. Figure 3.2 below presents a basic diagram of WRF and is from Skamarock (2008).

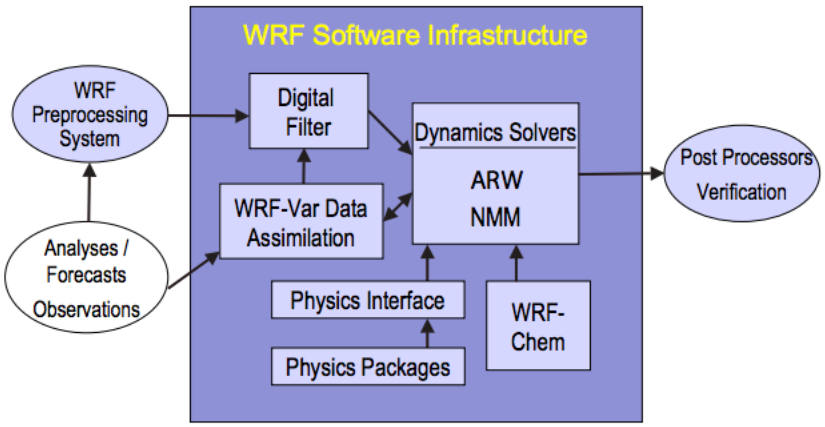

Figure 3.2: WRF System from Skamarock (2008). WRF obtains its ICs and BCs from a regional model and then generates a forecast.

There are two dynamics solvers for WRF, the non-hydrostatic mesoscale model (NMM) and the ARW. NMM and ARW both encompass physics schemes, numerics/dynamic options, initialization routines, and a data assimilation package (MM5 3-DVAR and WRF-Var) to produce simulations. ARW is used in this study, and the details to follow refer to this package. The governing equations are compressible, non-hydrostatic flux-form Euler equations and are written in sigma coordinates to address atmospheric variables, map projections, the coriolis effect, and the curvature of the earth. Perturbation forms of the governing equations can be found in Skamarock (2008).

There are many different options for running WRF, so it can be hard to decide which ones should be used. Planetary boundary layer (PBL) schemes have been widely

studied, but no single scheme stands out as the best, depending on the application. Three PBL schemes were studied over southeast TX with WRF Version 3.01 and the simulations found that the Yonsei University (YSU) scheme (Hong et al., 2006) and the asymmetric convective model (ACM2) (Pleim, 2007a,b) both were robust schemes, and the Mellor-Yamada-Janjic (MYJ) scheme (Mellor and Yamada, 1982) had the most bias in temperature and moisture (Hu et al., 2010). Coniglio et al. (2013) studied convection-allowing WRF forecasts with different PBL schemes and found that the YSU and ACM2 performed well, but the Mellor-Yamada-Nakanishi-Niino (MYNN) scheme (Nakanishi and Niino, 2009, 2004) was nearly unbiased in PBL depth, moisture, and potential temperature. Given that convection was studied here, at least in most simulations, the MYNN scheme will be used for the PBL.

Multiple microphysics schemes are available in WRF for dealing with precipitation and clouds. Otkin et al. (2006) demonstrated that based on the scheme, precipitation can vary significantly, so the microphysics scheme is held constant in all simulations. Following Coniglio et al. (2013), Thompson microphysics (Thompson et al., 2008) and the Noah Land Surface Model (Noah LSM) (Chen and Dudhia, 2001; Ek et al., 2003) were utilized in this study. For radiation, both short and long wave radiation, the Rapid Radiative Transfer Model with trace Gases (RRTMG) was used (Iacono et al., 2008). Table 3.2 below lists the options that were selected for the WRF simulations. Since the grid spacing is 3 km, convection is not parameterized, but is explicity resolved and calculated.

Table 3.2: WRF: Model configuration

WRF Version	3.6
Domain	1
Grid Size	425 X 425
Model Top	50 hPa
Vertical Levels	40
Grid Spacing	3 km
Time Step	18 sec.
PBL Scheme	MYNN
Land Surface Model	Noah
Long-wave radiation	RRTMG
Short-wave radiation	RRTMG
Microphysics	Thompson
Cumulus Scheme	None

3.5 WRF-DA Details

WRF-DA has a 3D-Var package and was used for this study. Figure 3.3 below shows the relationship between WRF-Var, datasets, and other components of NWP (Skamarock, 2008). WRF-DA, like any DA system, needs three basic pieces of information, a first guess of the atmosphere ($\boldsymbol{x_b}$), observations along with the error covariance matrix ($\boldsymbol{y}$ and $\boldsymbol{R}$), and a background error covariance matrix ($\boldsymbol{B}$). A first guess is obtained with NAM IC's in the WRF. There are certain errors and biases associated with the first guess, and a good estimation of $\boldsymbol{B}$ is essential for good results (Navon, 2009; Kalnay, 2003). Furthermore, forecast error covariance estimation for

the specific domain is necessary for quality results (Barker et al., 2012). The NMC method for estimating $\boldsymbol{B}$ is done by using the gen_be utility in the WRF-DA package. This takes the difference of forecasts that have a length of 12 and 24 hours and are valid at the same time to generate a background error. Sixty 24-hour simulations were used from 14 March 2015 to 12 April 2015 to estimate $\boldsymbol{B}$. Observation errors from the U.S. Air Force Weather Agency are defaults in WRFDA and are used here (Barker et al., 2012).

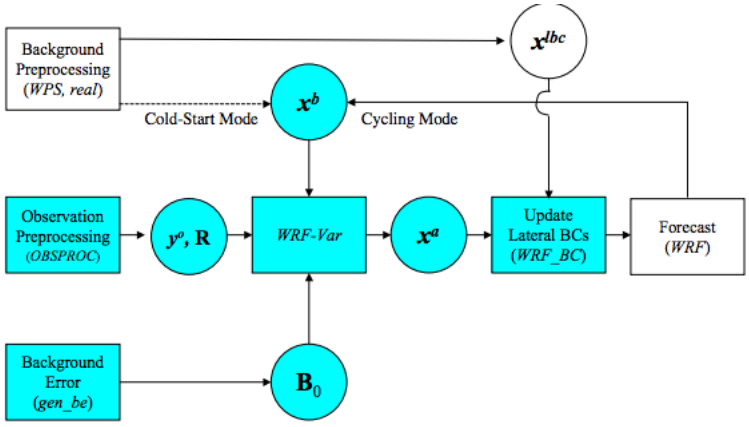

Figure 3.3: WRFDA System (Barker et al., 2012)

Quality control within a DA system is important and WRF-Var has procedures to reject bad data. Observations are rejected if the differences between the observations and first guess are larger than a specified threshold. This can vary based on errors of the background error and the observation error. This can be problematic if the first guess is a poor forecast. In this case, some good observations are rejected. In WRF-DA version 3, an outer loop is included to alleviate this effect. If observations are rejected at the initial assimilation time, they have a chance of being accepted later. The updated guess will become the first guess, allowing the comparisons between

the rejected observations and the first guess to take place again. Using two to three outer loops is common; three loops are used in this study.

3.6 Research Design

WRF simulations with different DA strategies were done to test the sensitivity of Texas A&M RAOBs in a forecast. Spin-up time in a NWP model helps microsphysical parameters develop (Done et al., 2004) and allows realistic looking convection to start taking place. All simulations have three hours of spin-up time before the official start time and final run of 24 hours.

The first simulation will be a WRF only case (WRF_C), with ICs and BCs from the NAM-12 km model. This means that for an official start time of 18Z, the simulation is spin-up starting at 15Z, using a three hour forecast from the NAM, and was run for three hours, before the 24-hr forecast period begins at 18Z. The other simulations are slight modifications of this case. Below is Figure 3.4, which shows the setup of the model cycling. WRF_R will be the second simulation and will spin-up as the first run and only assimilate the RAOB at time t. WRF_OB assimilated the MADIS observations mentioned earlier starting at time t-3, and then assimilated data at 1-hour intervals through time t. WRF_OBR follows the same strategy as WRF_OB, but also incorporates the RAOB at time t, along with the other observations. Simulations were done using accounts through the supercomputing center at Texas A&M University.

Verification of model performance was done using Stage-IV precipitation data, MADIS observations (Miller et al., 2005a), the RTMA (Pondeca et al., 2011), and the RAP operational model, which replaced the Rapid Update Cycle (RUC) (Benjamin et al., 2004) in 2012. The Developmental Testbed Center (DTC) has tools for evaluating model performance called Model Evaluation Tools (MET) (Brown

et al., 2009), and they were used extensively, along with SPoRT-MET (Zavodsky et al., 2014), which helps run certain parts of MET. Predicted variables and derived products of interest include: surface temperature, surface dewpoint, surface wind, upper-level temperature, upper-level moisture, upper-level wind, radar reflectivity, and accumulated precipitation. Statistics and characterizations for the winter and convective cases were done.

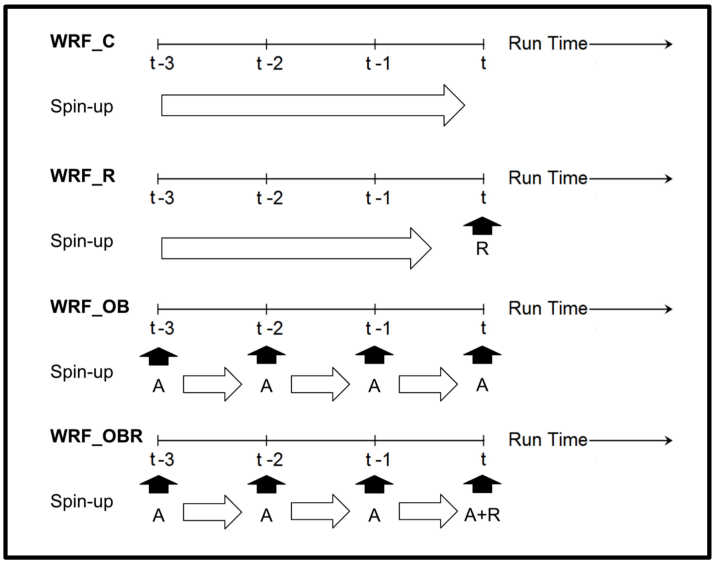

Figure 3.4: DA strategies. WRF_C is the control case that does no DA, WRF_R only assimilates at the RAOB valid time (designated by R), WRF_OB assimilates MADIS data at hour intervals starting at t-3 and every successive hour (A), and WRF_OBR assimiliates the same data as the previous, but with the RAOB included at time t (A+R).

4. RESULTS

Forecast verification determines the quality of forecasts. This can be done in two ways: subjectively or objectively. Subjective analyses are based on a human judgement and therefore do not always take into account quantities in a forecast. They are often useful, however, because a human judges forecasts differently than a computer. A forecast may, for example, do a poor job predicting temperature with a cold front passage, but the timing might be correct. Objective analyses, on the other hand, is based more on predicted quantities and can remove possible biases that a person might have. An objective analysis is also much faster than a subjective one, making it an easy choice. Objective verification will be used for all simulations, and in addition, subjective verification will be done for a few cases. Two areas of verification are used, one for the whole model domain, and one for a smaller boxed region around the RAOB launch site at Texas A&M (Figure 4.1). This smaller region was chosen in order to capture the local influences of the RAOB data at Texas A&M and downstream.

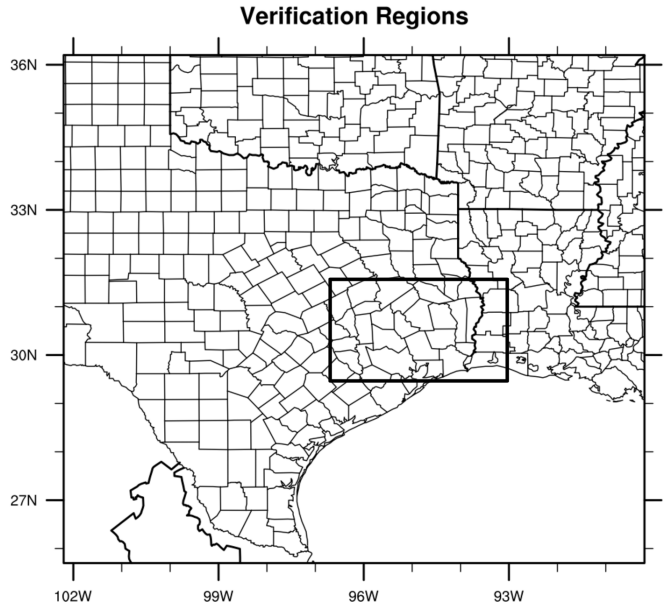

Figure 4.1: Two verification regions for the model. The smaller region is positioned over College Station, TX and downstream to capture local affects of the RAOB.

4.1 Objective Precipitation Verification

Precipitation is often verified objectively using derived precipitation variables from simulations and mapping them to the radar observation data. Counts of forecast-observation pairs are computed to complete a contingency table (Wilks, 2011). Figure 4.2 shows this contingency table. Typically, it is used for precipitation and severe weather verification. The table is not limited to this verification, but it is very useful in in regards to precipitation. There are many metrics derived from this table, but this study will only use several that are often used for precipitation verification. Among the ones used here are probability of detection (POD), false alarm rate (FAR), critical success index (CSI), which is also known as the threat score (TS).

Several quantities used here deserve explanation. POD = $\frac{a}{a+c}$ and is the fraction

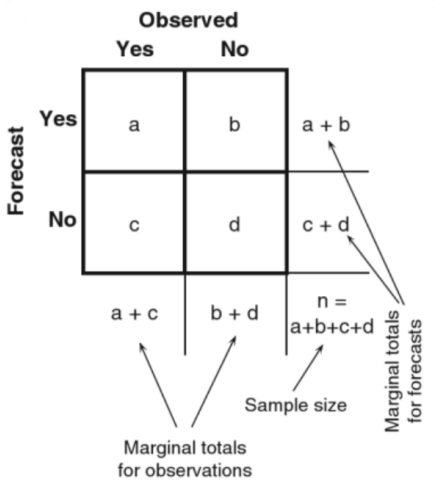

Figure 4.2: Contingency table for forecast verification from Wilks (2011). The a forecast-observation pairs are hits, the b pairs are false alarms, the c pairs are misses, and the d pairs are correct rejections.

of times a forecast event occurred when forecasted, FAR = $\frac{b}{a+b}$ and is the fraction of yes forecasts that are wrong, and CSI = $\frac{a}{a+b+c}$, which is the number of correct forecasts divided by the total number of hits, misses, and false alarms. An ideal precipitation forecast would have POD = 1, FAR = 0, and CSI = 1. In addition to using these scores, different precipitation quantities can be verified. This section focuses on accumulated precipitation, but radar reflectivity is compared to simulated reflectivity later. Different thresholds of precipitation accumulation can be verified with these metrics. For the convective cases, thresholds of 25 mm, 10 mm, 5 mm, and 1 mm were used, while the winter cases had lower thresholds of 5 mm, 3 mm, and 1 mm due to lower precipitation amounts. For both cases, three hour accumulations were used.

As mentioned earlier, the SPoRT-MET (Zavodsky et al., 2014) verification tools were used extensively for verification. For the precipitation verification, Stage IV radar data was compared to the model accumulated precipitation. This radar data

is on a 4 km grid, so simulation data at 3 km is mapped to this data. If grid boxes of model and observations were directly compared, skill scores would be low due to positioning and timing of rainfall. Neighborhood techniques, however, reward close forecasts; this approach is used. The neighborhood box for these simulations was 9 grid points. Since the radar data resolution is 4 km, the verification is done on a 36 km resolution scale for the whole domain. An example is necessary to explain the neighborhood verification. 25 mm of accumulated rainfall is observed by radar for a given three hour period, and if the model forecasted this within the 36 km grid box, this is a hit.

4.1.1 Convective Precipitation Cases

The convective weather situations are verified first, and for the whole domain of the model in Figure 4.1. CSI, POD, and FAR for 1 mm, 10 mm, and 25 mm are below in Figures 4.3 — 4.5. 5 mm verification is not pictured here since it is almost identical to Figure 4.3 For CSI, WRF_OB and WRF_OBR performed slightly better than the other simulations at early forecast hours, except at 25 mm. These differences are attributed to the generally better performance in POD during these hours and equal performance in FAR. WRF_R and WRF_C both were very similar for all thresholds. WRF_R, however, has some minor improvements at the 25 mm threshold. POD is better at later forecast hours, which subsequently gives it a slight bump in CSI at these hours. FAR for all cases and thresholds is close to identical. FAR increases with time as the model simulation quality degrades. It should be generally noted that as thresholds are increased from 1 mm to 25 mm, CSI and POD decrease, while FAR increases.

The small domain for verification, seen in Figure 4.1, is verified next for the convective cases. This domain was chosen because it is close to and downstream of

the RAOB launch location. Figures 4.6 - 4.8 show the three variables for the same precipitation thresholds. The CSI for WRF_OB and WRF_OBR is higher than the others in the beginning of the forecast period, but even more so than for the whole domain verification. The precipitation threshold of 25 mm is different than the other graphs. CSI and POD for the WRF_R simulation is much higher than the others at the later hours in the forecast. POD in Figure 4.5b has this feature, but it is not as noticeable. CSI values of WRF_R are above the upper standard deviation bound of WRF_C for hours 21 and 22, but not for any other hours.

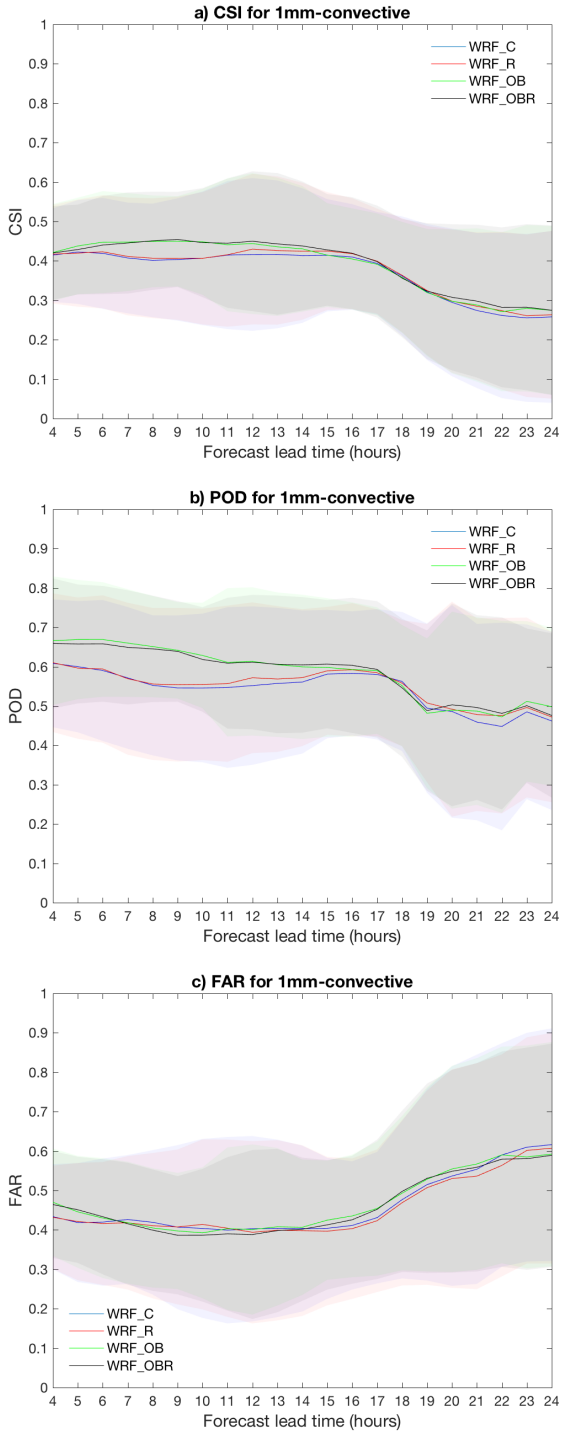

Figure 4.3: CSI, POD, and FAR for the 23 convective cases and the 1 mm precipitation threshold, whole domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

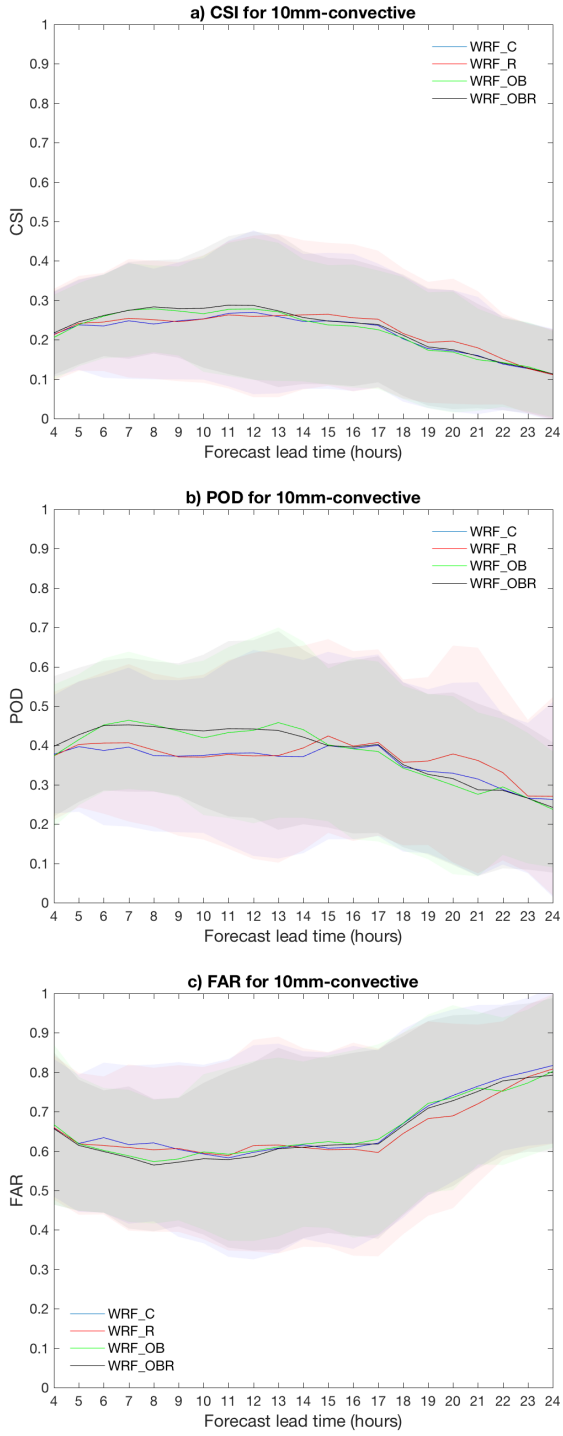

Figure 4.4: CSI, POD, and FAR for the 23 convective cases and the 10 mm precipitation threshold, whole domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

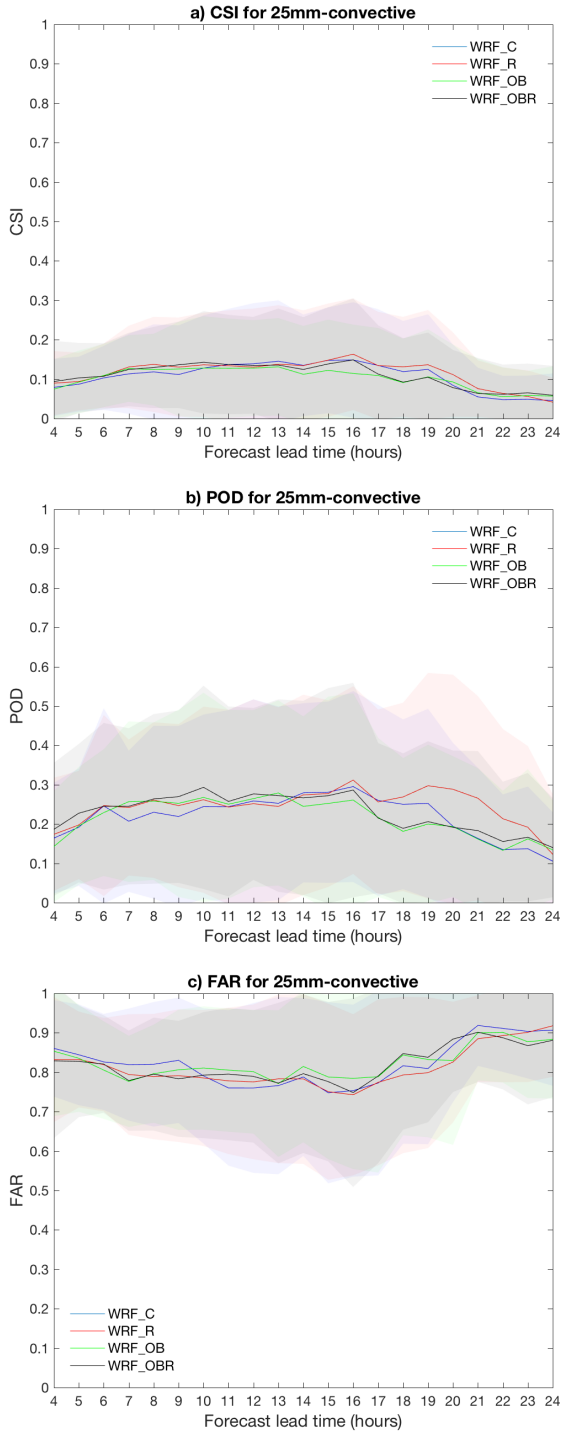

Figure 4.5: CSI, POD, and FAR for the 23 convective cases and the 25 mm precipitation threshold, whole domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

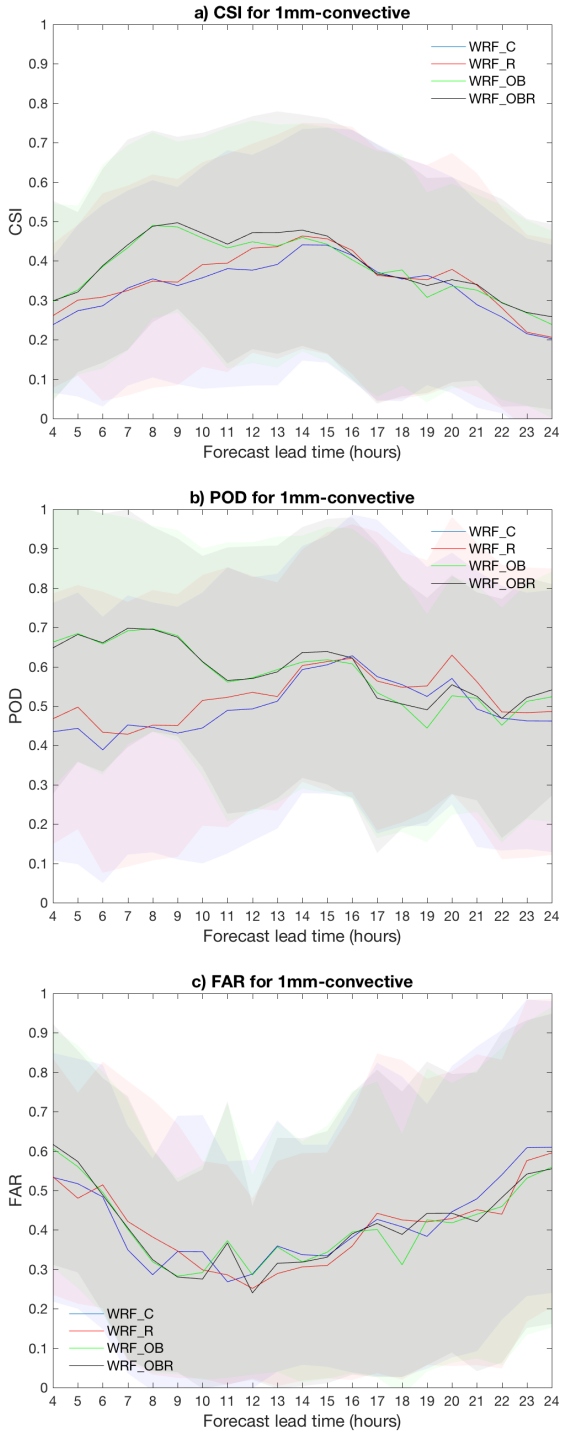

Figure 4.6: CSI, POD, and FAR for the 23 convective cases and the 1 mm precipitation threshold, small domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

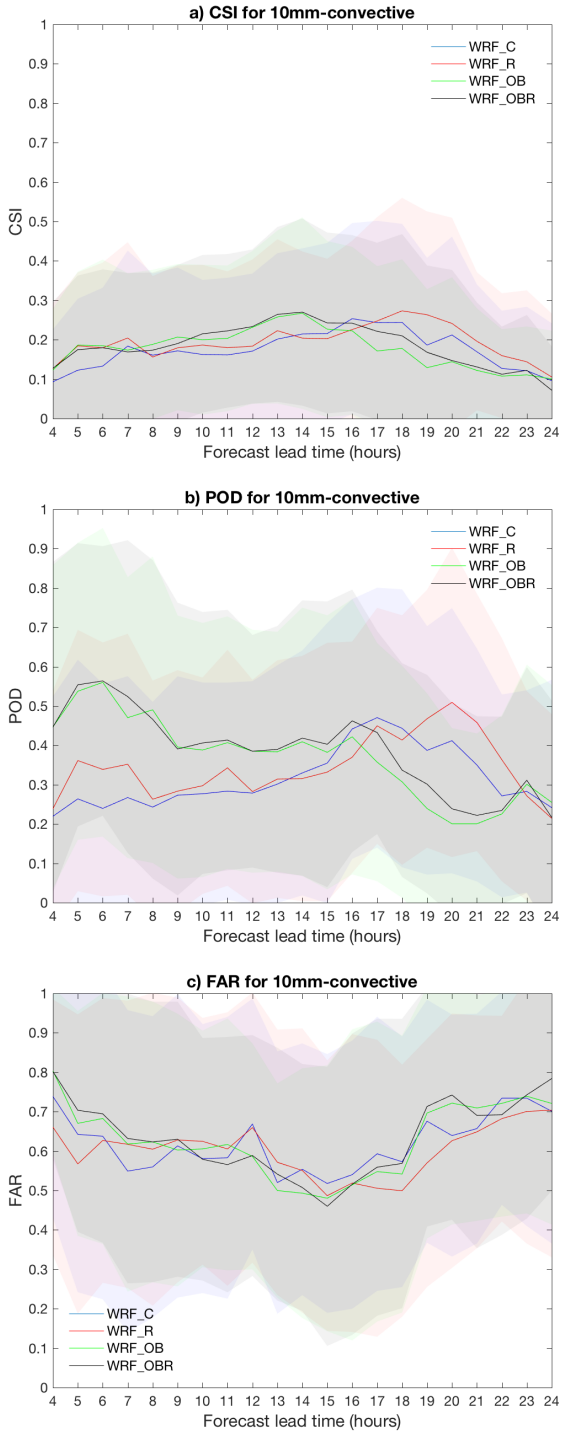

Figure 4.7: CSI, POD, and FAR for the 23 convective cases and the 10 mm precipitation threshold, small domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

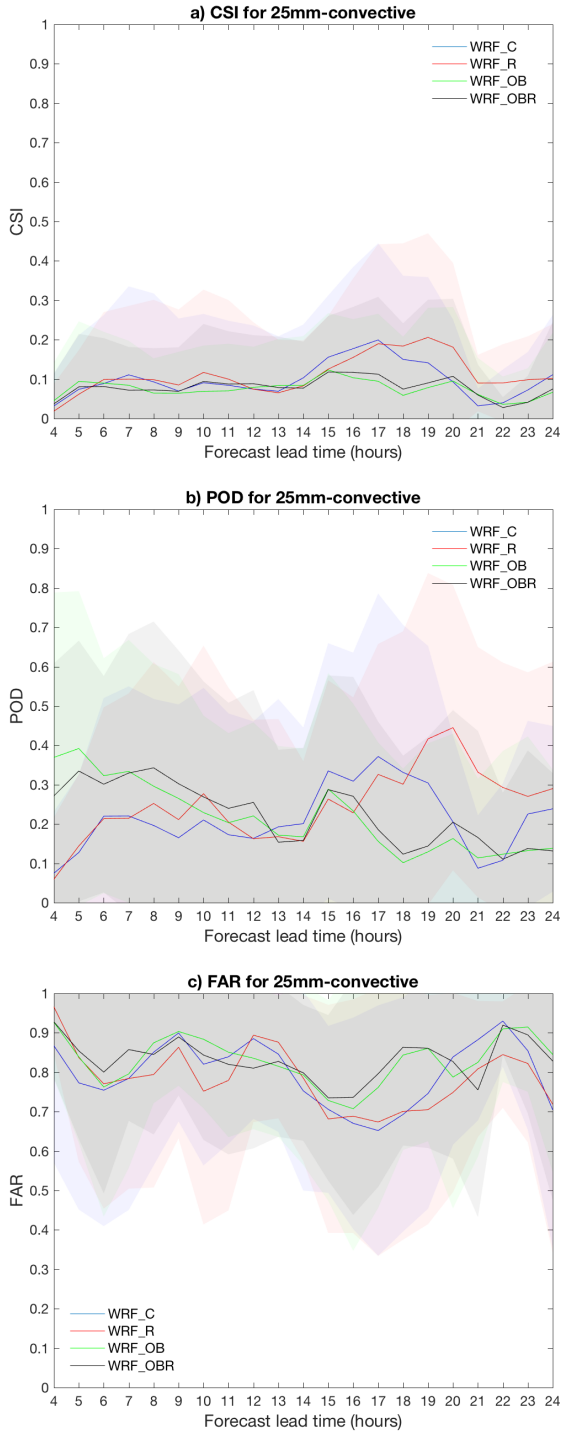

Figure 4.8: CSI, POD, and FAR for the 23 convective cases and the 25 mm precipitation threshold, small domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

4.1.2 Winter Precipitation Cases

The winter cases are now investigated. The verifications for 1, 3, and 5 mm are seen in Figures 4.9 — 4.11. Overall, the WRF_C and WRF_R simulations perform similarly and better than the other two simulations for CSI at all of the thresholds. Even though these perform somewhat better, there is substantial overlap of values at all thresholds. POD for the WRF_C case and 5 mm stands out as it is higher than the others from hours 8 to 10. Since there are only six cases for the winter, standard deviations are higher for the three variables at each precipitation threshold.

The smaller domain from Figure 4.1 is verified and not much different than the previous domain. Only the 1 and 3 mm precipitation thresholds had enough data to display meaningful graphs because of low precpitation amounts, which renders the standard deviations less useful because they include six or fewer data points. For the 1 and 3 mm thresholds (Figures 4.12 and 4.13, respectively), there is not a positive impact from any data assimilation scheme, except perhaps for 3 mm. A small positive impact in CSI is evident for the WRF_OB and WRF_OBR late in the simulation period. The WRF_R simulation scores are very similar to the rest of the simulations, with little difference observed for either threshold.

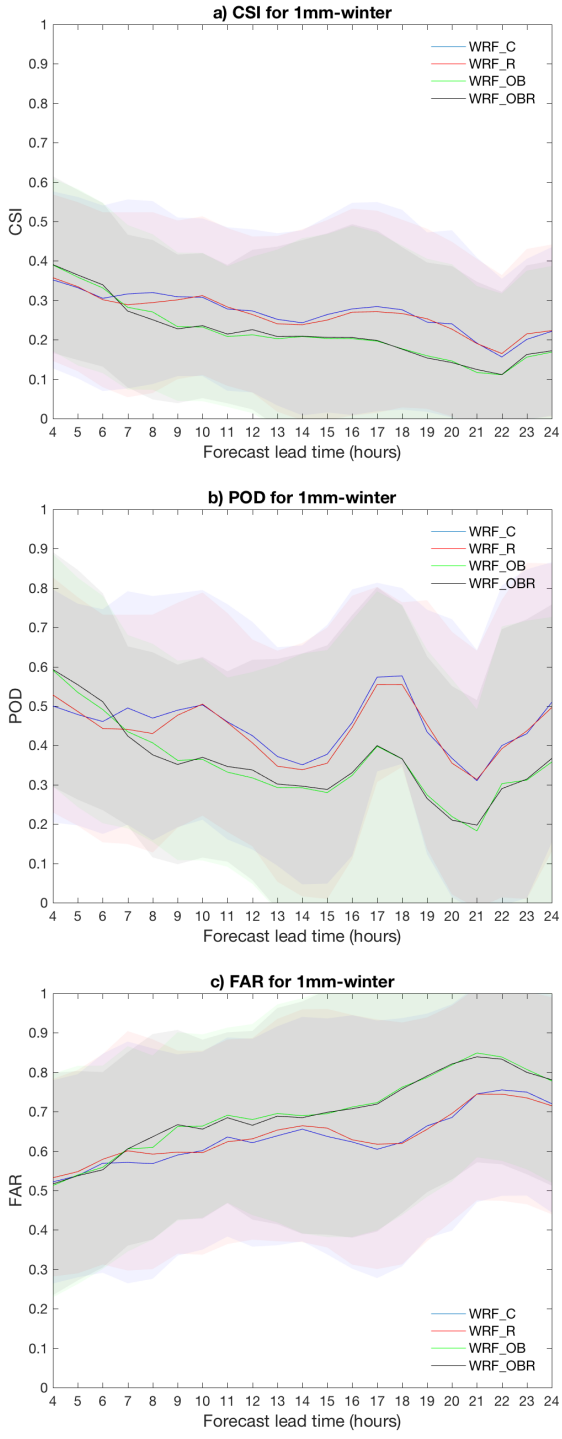

Figure 4.9: CSI, POD, and FAR for the six winter cases and the 1 mm precipitation threshold, whole domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

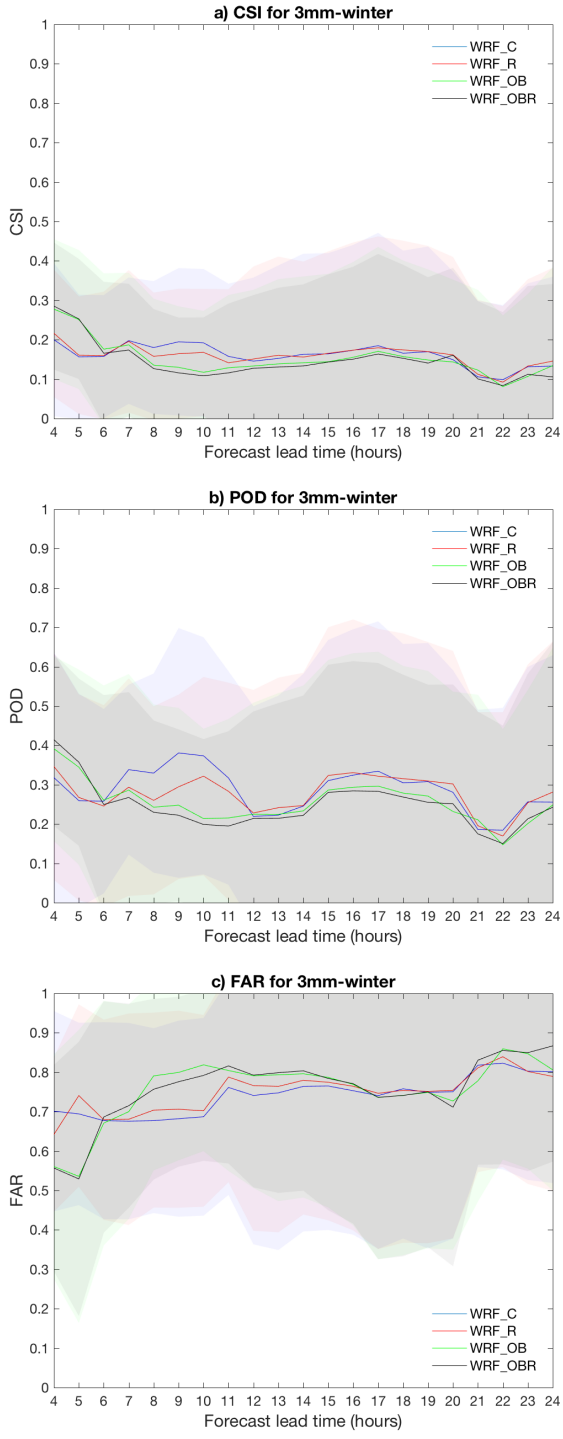

Figure 4.10: CSI, POD, and FAR for the six winter cases and the 3 mm precipitation threshold, whole domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

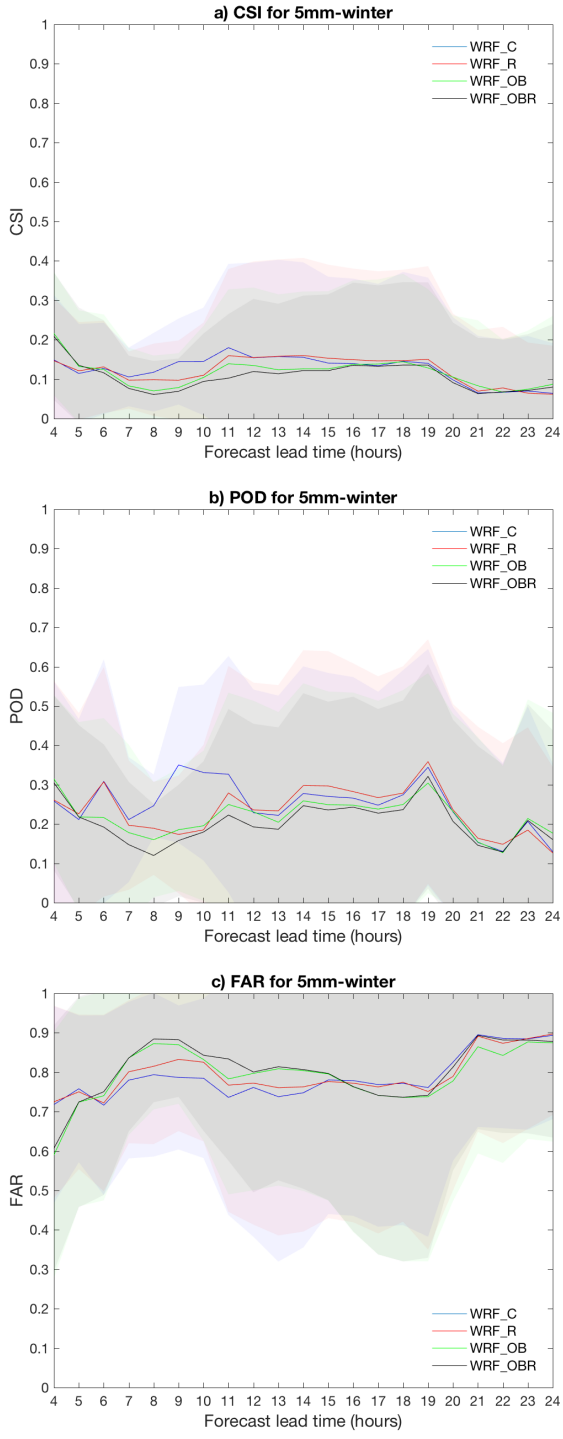

Figure 4.11: CSI, POD, and FAR for the six winter cases and the 5 mm precipitation threshold, whole domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

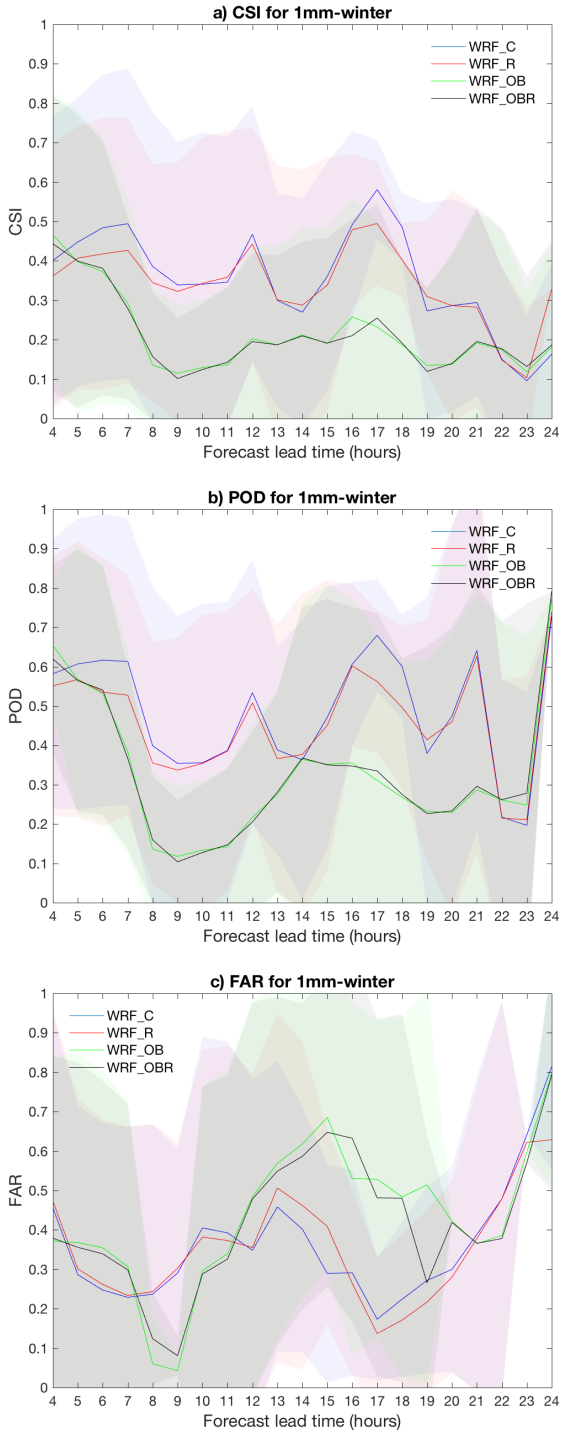

Figure 4.12: CSI, POD, and FAR for the six winter cases and the 1 mm precipitation threshold, small domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

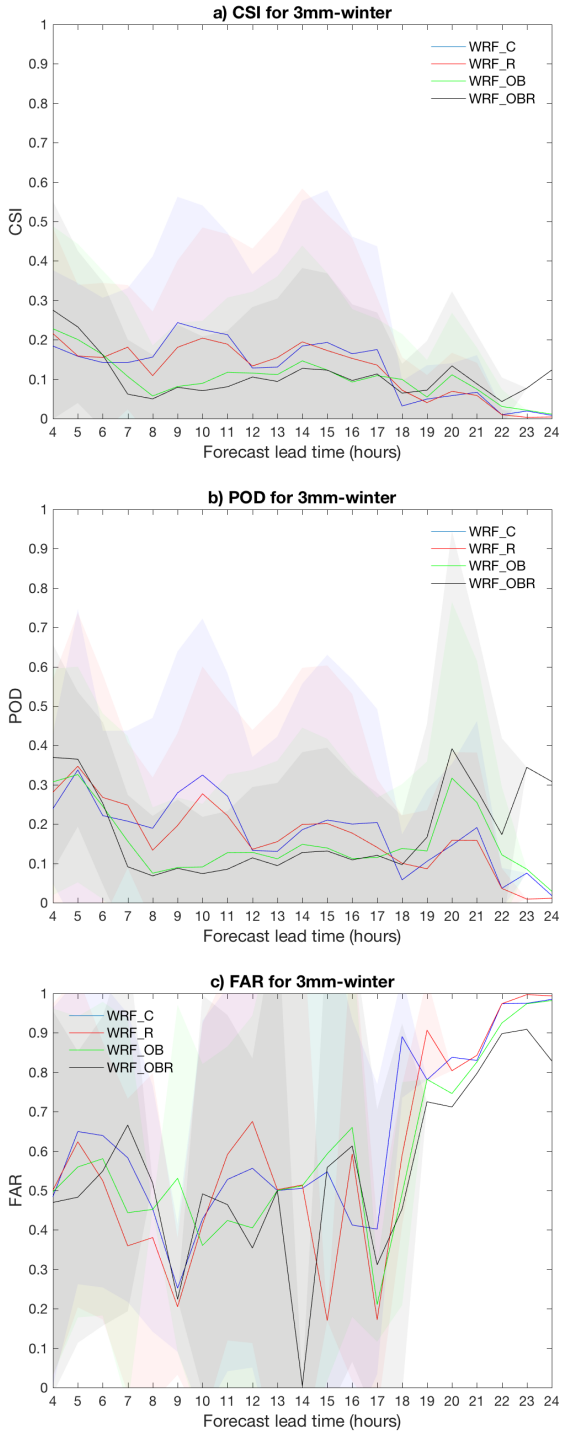

Figure 4.13: CSI, POD, and FAR for the six winter cases and the 3 mm precipitation threshold, small domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

4.1.3 Precipitation Verification Summary

DA impact on precipitation forecasting is limited. Generally speaking for the convective weather, WRF_OB and WRF_OBR produce the most improvement on precipitation forecasting overall. The WRF_R simulations assimilate a single RAOB, but have the single greatest positive impact on high precipitation thresholds. For forecast hours after 19, and the smaller verification domain, it performs better than the other simulations with a precipitation threshold of 25 mm. The assimilation of MADIS observations in WRF_OB appears to dampen the effect of RAOB DA on WRF_OBR, as little difference is observed between the two simulations. The winter cases are strongly affected by the DA as WRF_OB and WRF_OBR scores were significantly lower than WRF_C and WRF_R. WRF_R did not improve the scores either, with similar scores to WRF_C. Looking at the cases individually, only one DA case for WRF_OB and WRF_OBR had similar verification scores to WRF_C.

4.2 Verification of Important Convective and Winter Variables

When verifying atmospheric variables, mean squared error (MSE) is commonly used to measure accuracy. The equation below describes MSE

$$\frac{1}{M}\sum_{m=1}^{M}(y_m - o_m)^2 \tag{4.1}$$

where the squared differences are averaged over each pair of forecasts and observations. y_m represents the forecasted quantity at a location, while o_m represents the observation at that same location. More often, root mean squared error (RMSE) is used to verify forecasts. This is simply the square root of MSE and retains the same units as the verified variable. Temperature, dewpoint, and wind are verified at the surface using RMSE, while other levels at 500 mb, 850 mb, and 925 mb are presented in the form

of maps. As in the precipitation verification, convective and winter cases will be separate and verified with two separate domains seen in Figure 4.1.

Surface variable verification can use many data sources from MADIS for point verification: metars, maritime, meso-nets, and SAOs. Because of the poor quality of some meso-nets, they are not used to verify the model data here. Another dataset used for surface verification is the Real-Time Mesoscale Analysis, which combines model output and observations to produce a high-resolution analysis at the surface. With a resolution of 2.5 km, it is very close to the resolution of the WRF simulations. For certain verifications, spatial maps of error will be plotted using this dataset.

Upper-level verification often takes place with observations, but model analyses are also used. The RAP, as mentioned in Section I, replaced the RUC in 2012 and is a 13 km operational model for North America. It has a new analysis every hour, 40 vertical levels, extends to 10 hPa, and assimilates more data than its predecessor. This model is used for verification of temperature and relative humidity (RH) at 500, 850, and 925 mb.

Maps showing spatial impact of the different DA simulations will be plotted in addition to time-series plots in this section. Since this study hypothesizes that there is some benefit from assimilating the RAOB launched at Texas A&M University, a map displaying a benefit from the RAOB assimilation can be plotted. This will help in determining the radius of influence and spatial impact that the RAOB has on the simulation. To do this, absolute error for each simulation must be calculated (WRF_C, WRF_R, WRF_OB, and WRF_OBR) with respect to the RTMA at the surface or the RAP at upper-levels for the interest variable. The benefit maps will be for WRF_R and WRF_OBR since they assimilate the RAOB. Once the absolute errors are calculated, the maps are made for WRF_R and WRF_OBR benefit respectively. For the first map, WRF_R values are subtracted from WRF_C values (WRF_C minus

WRF_R), and for the second, WRF_OB values are subtracted from WRF_OBR values (WRF_OB minus WRF_OBR). The result is positive numbers when the forecast better matches the analysis, and negative when it does not match the analysis as well. This improvement will be called benefit.

4.2.1 Convective Cases

Figure 4.14 below is the temperature RMSE for the convective cases over the whole verification domain. WRF_R has the highest RMSE at start time, WRF_C has a slightly lower value, while WRF_OB and WRF_OBR have the lowest values. The average WRF_OB and WRF_OBR values are lower than the other two cases by almost 0.5 C at the beginning of the simulation, but they soon join all simulations and agree within tenths of a degree. Since the WRF_R simulation only assimilates RAOB, it is not surprising that its influence is marginal.

The next verification variables are surface dewpoint and wind speed in Figure 4.15 and Figure 4.16. The dewpoint errors have small differences between the four cases with WRF_OB and WRF_OBR having somewhat lower errors at the beginning of the simulation period. Wind speed verification shows no noticeable difference between the simulations.

The smaller domain in Figure 4.1 is verified now with surface temperature. RMSE is seen in Figure 4.17, with small differences in the average RMSEs. Overlap in standard deviation also indicates that there is low statistical significance. RMSE errors for surface dewpoint and wind were not included because there was nothing significant about any simulation.

As suggested by Figure 4.17, the graphs of the spatial benefit for WRF_R and WRF_OBR at the surface would not be interesting, so it is not shown. Two upper-level verifications are more interesting. Early in the convective forecast at 3 hours,

benefit from WRF_R and WRF_OBR is apparent in 850 mb RH. Figure 4.18 shows slightly more positive benefit from the WRF_R simulation. The benefit for both simulations is effectively gone several hours later. This larger benefit is expected for WRF_R, as the only observation assimilated here is a RAOB. WRF_OB assimilates many observations, so impact of the RAOB is masked in WRF_OBR.

Temperature benefit plots for 925mb are shown for F05 in Figure 4.19 and shows benefit that is migrating to the north. This would be expected for 925 mb for convective cases because winds would typically be to the north. WRF_OBR shows less benefit than WRF_R, but it is still apparent in the plots.

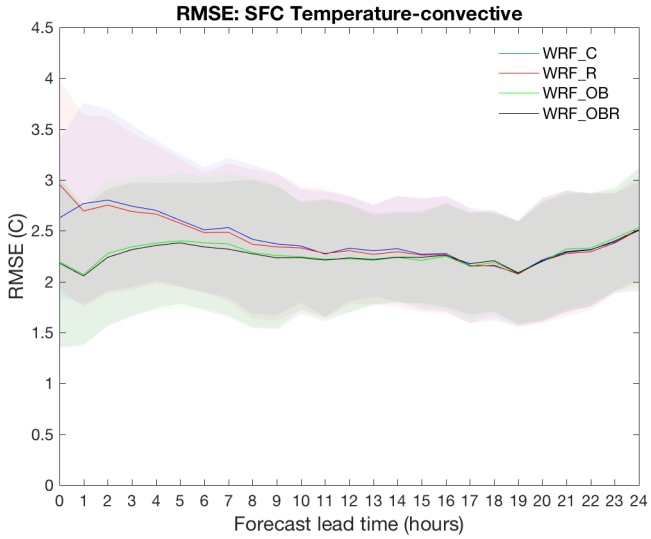

Figure 4.14: Surface temperature RMSEs for all 23 convective cases, whole domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

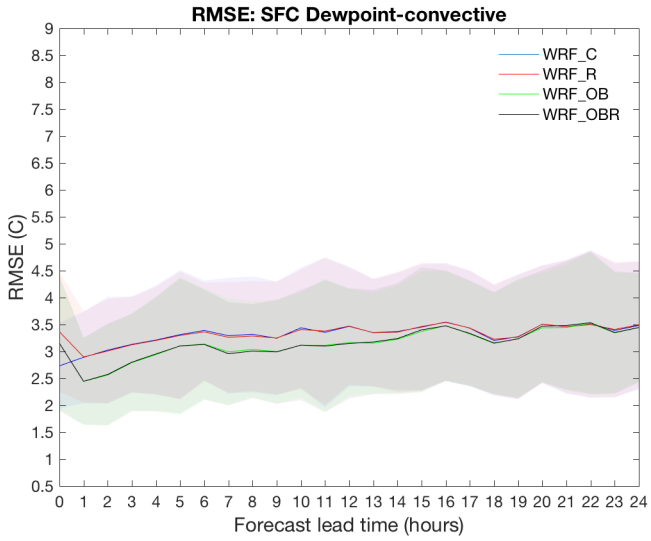

Figure 4.15: Surface dewpoint RMSEs for all 23 convective cases, whole domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

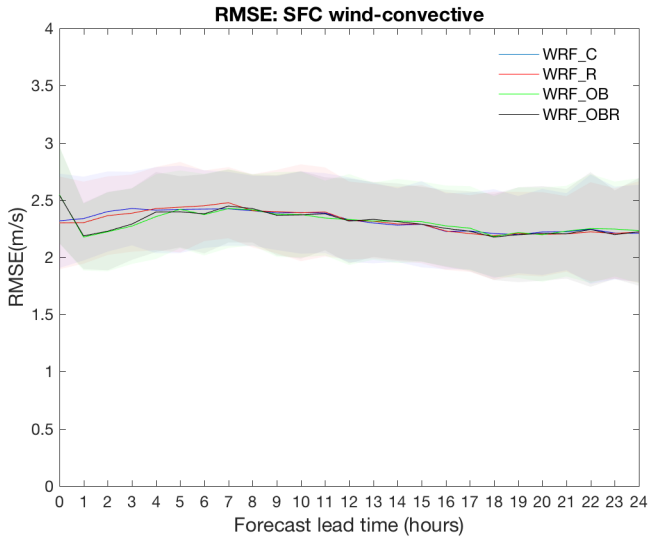

Figure 4.16: Surface wind RMSEs for all 23 convective cases, whole domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

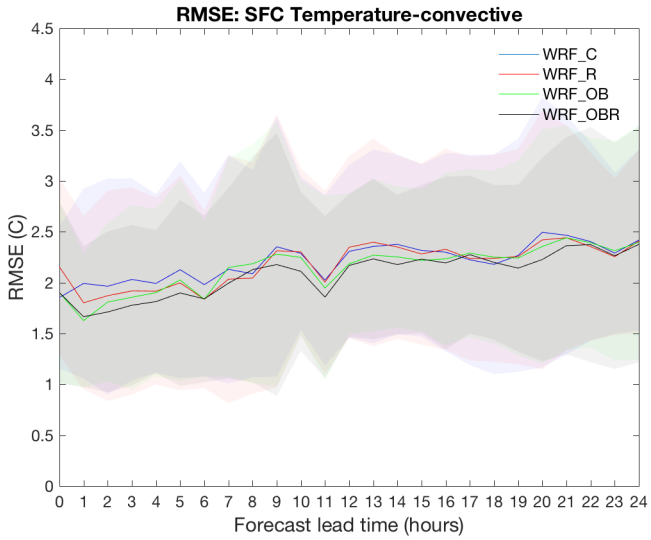

Figure 4.17: Surface temperature RMSEs for all 23 convective cases, small domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

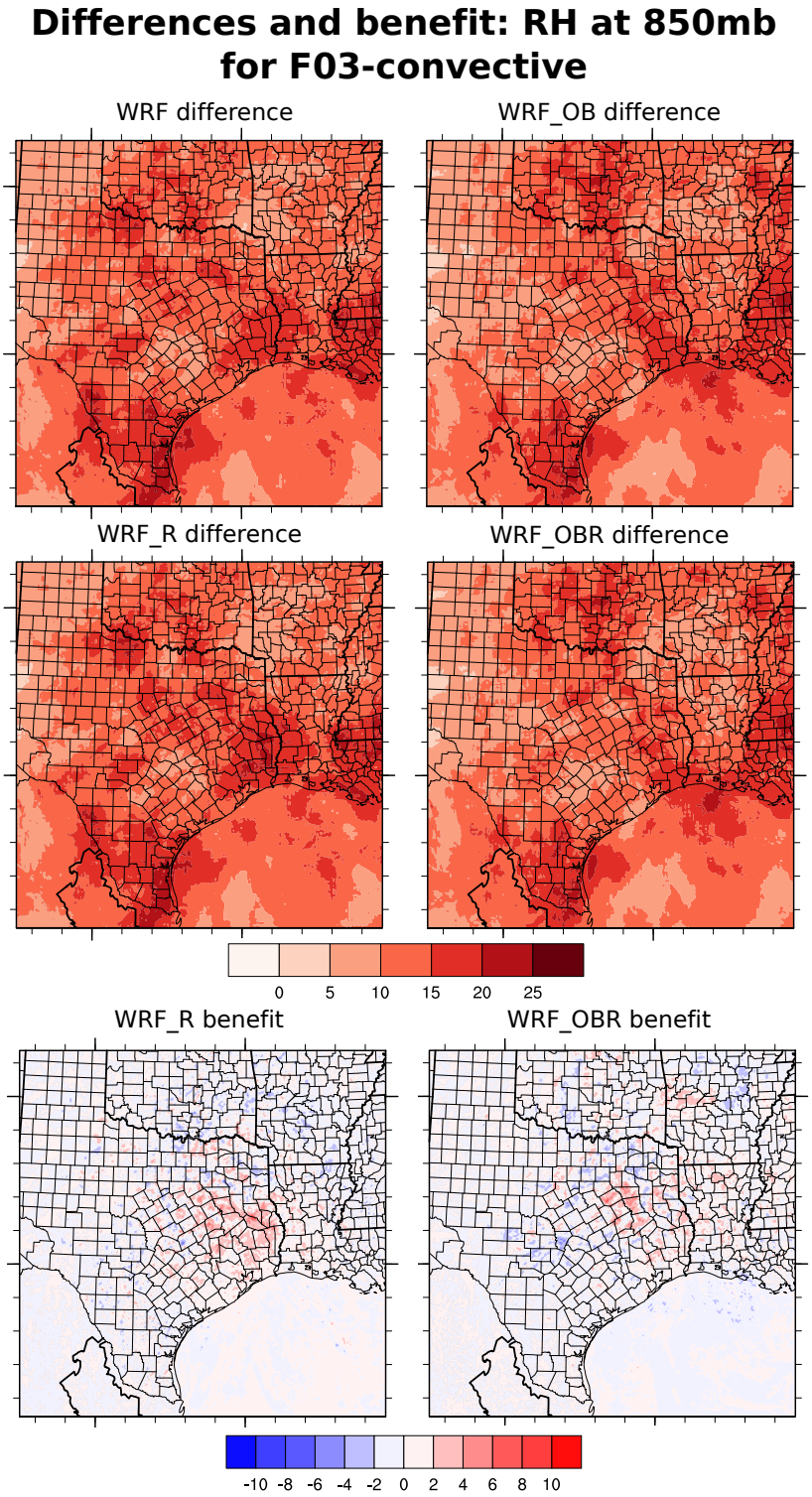

Figure 4.18: RH benefit (%) from the WRF_R and WRF_OBR simulations at 850 mb and at a lead time of 3 hours.

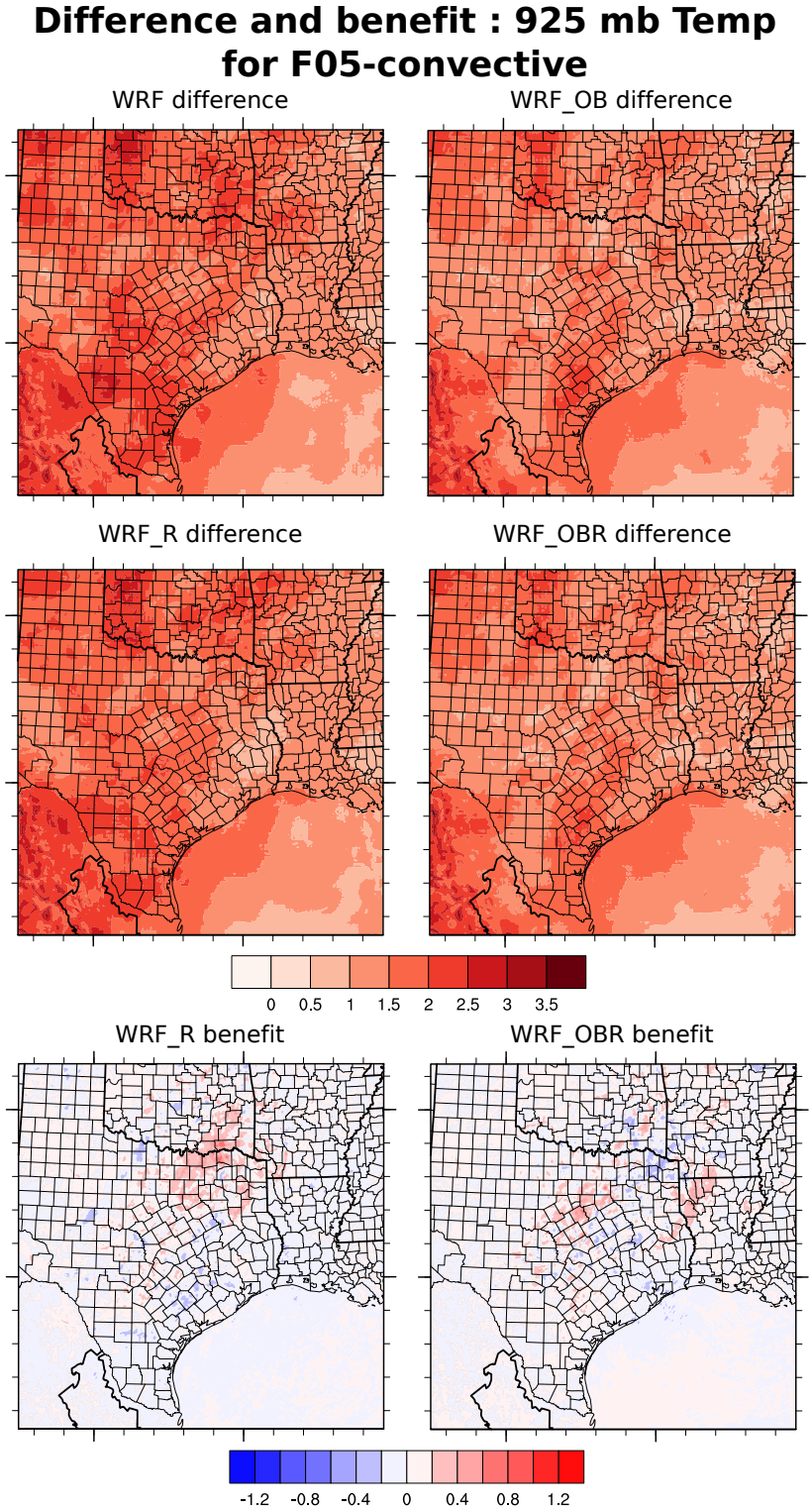

Figure 4.19: Temperature benefit (C) from the WRF_R and WRF_OBR simulations at 925 mb and at a lead time of 5 hours.

For the convective simulations, the WRF_OB and WRF_OBR simulation performed the best in terms of surface temperature and dewpoint RMSEs. WRF_R did not deviate much from WRF_C, which suggests that the RAOB had little impact on the forecasted surface temperature. Most significant for WRF_R and WRF_OBR is their influence on the short term forecast of RH at 850 mb and temperature at 925 mb. The spatial impact was not particularly impressive, but there was a positive signal in both spatial maps that assimilated the RAOB. Not only is RH in the boundary layer particularly important when forecasting convection, but 925 mb temperature is also important for forecasting convection. Benefit for the simulations was not incredibly significant when the RAOB was assimilated in the WRF_R and WRF_OBR case, but is highest for the WRF_R case. MADIS observations assimilated in WRF_OB limited the impact of the RAOB in the WRF_OBR case.

4.2.2 Winter Cases

The six winter cases were also verified with the same datasets over the two domains. Results for surface temperature and dewpoint are below in Figure 4.20 and Figure 4.21 for the whole domain. Verification for both variables is better at all hours for WRF_OB and WRF_OBR. This result is different from the convective cases as there was a small improvement for WRF_OB and WRF_OBR. In addition, standard deviation overlap is low at all forecast hours, indicating that these errors are statistically different than WRF_C and WRF_R. Surface wind verification is similar to the convective cases.

For the smaller area, Figure 4.22 shows surface temperature verification. Noticeable differences are apparent for the first nine hours, then all four simulations roughly agree. WRF_OB and WRF_OBR are again the best performers during the beginning of the simulation for surface temperature, along with WRF_R. Surface dewpoint shows

almost equal differences among the simulations and had large standard deviations for both WRF_OB and WRF_OBR.

When plotting the RTMA temperature benefit at 2 meters for the winter case, there is a noticeable benefit from 0 to about 12 hours near and around College Station in the spatial images for WRF_R. Figure 4.23 shows absolute temperature differences for simulations at F03, then the individual benefit of WRF_R and WRF_OBR. There are improvements near 1 C for the WRF_R simulation, but limited benefits for WRF_OBR over the Gulf of Mexico. The RAOB has a large benefit when it is the only observation assimilated, but benefit disappears when the other observations are assimilated.

Anther level important for winter weather temperature besides the surface is 925 mb because this is often the location of the elevated warm layer. Figure 4.24 shows the absolute differences at 925 mb of temperature for winter simulations at forecast hour five, and then the benefit for WRF_R and WRF_OBR. The benefit is more pronounced for WRF_R as in the previous figure, but it is nonetheless present in the WRF_OBR image too. This benefit is important because the location of the elevated warm layer is important for winter precpitaion type forecasting.

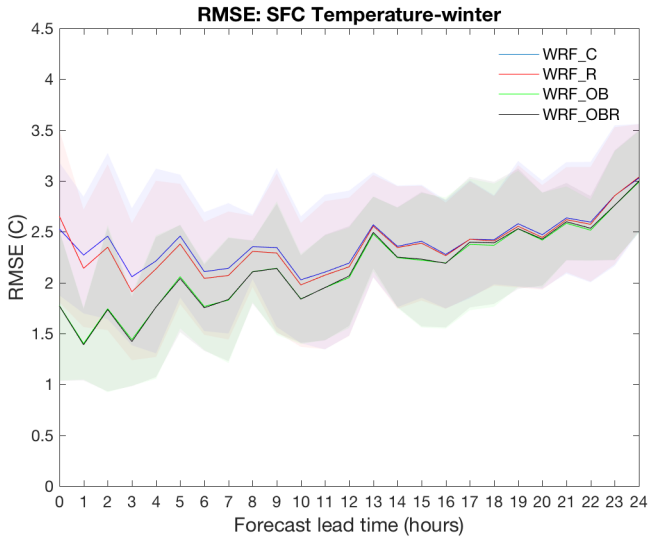

Figure 4.20: Surface temperature RMSEs for all 6 winter cases, whole domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

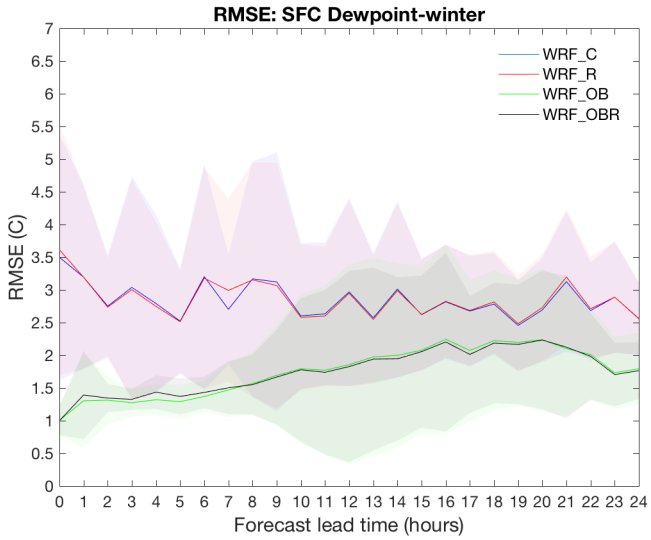

Figure 4.21: Surface dewpoint temperature RMSEs for all 6 winter cases, whole domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

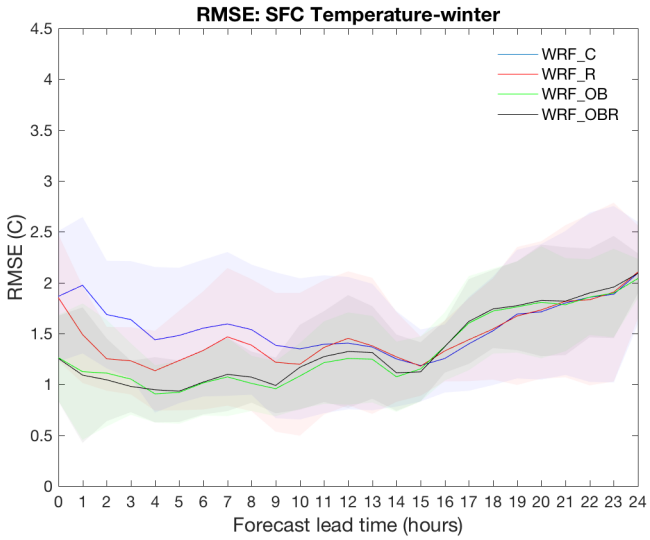

Figure 4.22: Surface temperature RMSEs for all 6 winter cases, small domain verification. The solid lines are the average scores for all cases, while the shading represents the standard deviation at the specified hour.

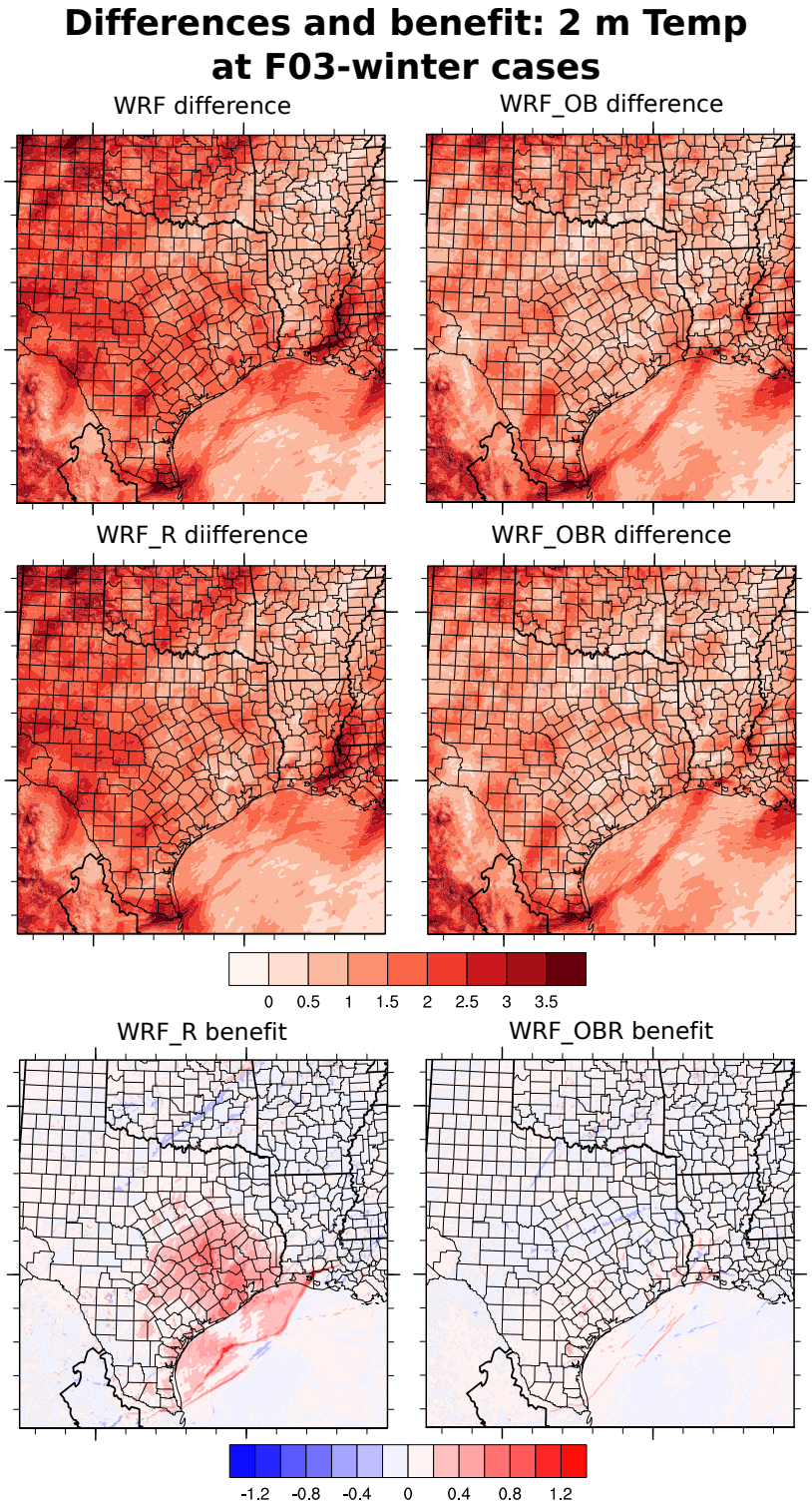

Figure 4.23: Absolute differences for each simulation for RTMA 2 meter temperature at F06 for winter cases. The other images are benefits for WRF_R and WRF_OBR.

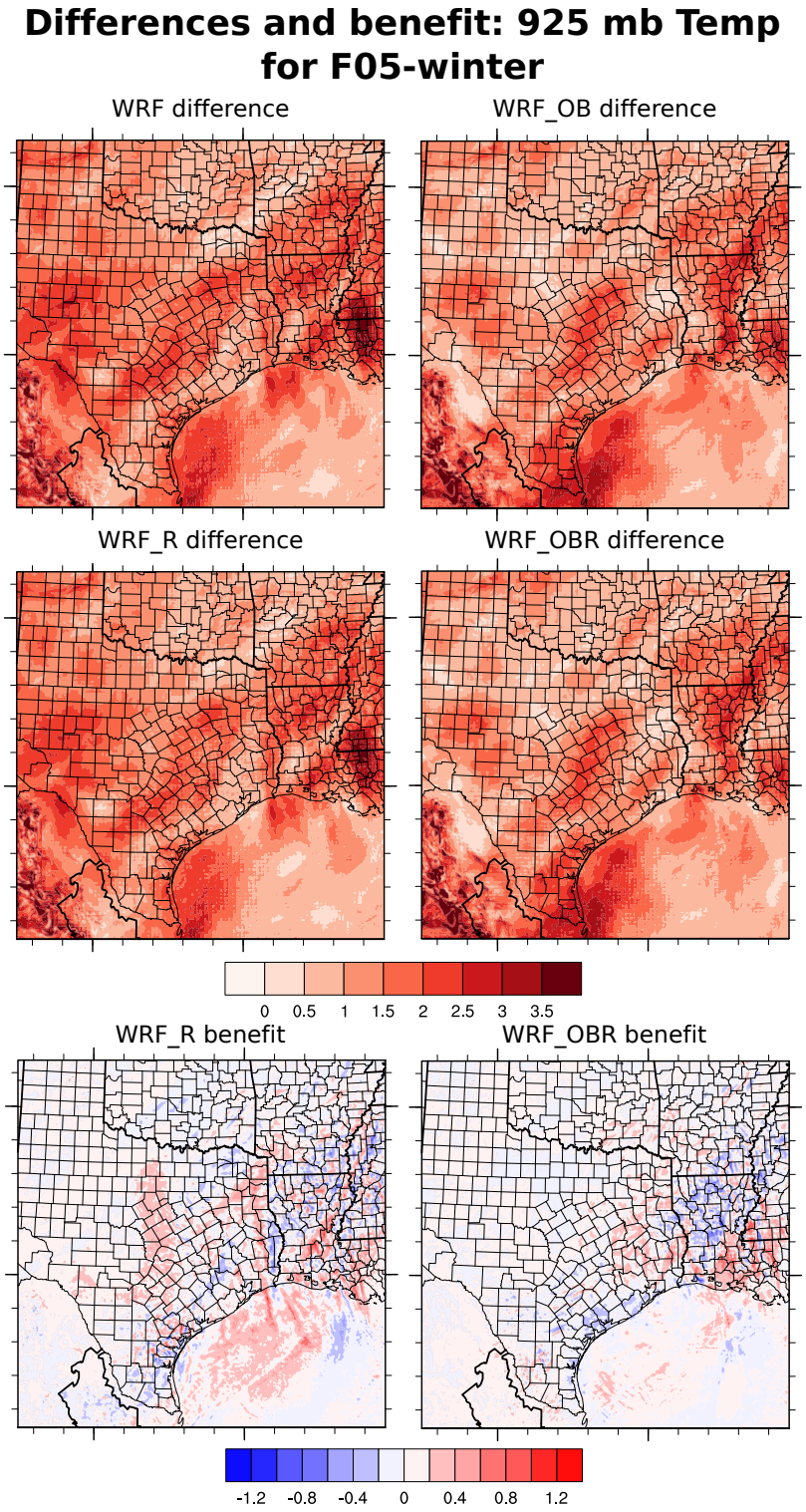

Figure 4.24: Absolute differences for each simulation for RAP 925 mb temperature at F05 for winter cases. The other images are benefits for WRF_R and WRF_OBR.

WRF_OB and WRF_OBR are clearly the best performers for the winter simulations, as they had much lower RMSEs for surface temperature, surface dewpoint, and had lower absolute values on the spatial upper-level maps. This seems unusual since they performed poorly for precipitation verification. WRF_R alone, however, did improve the surface and low-level temperatures significantly for about the first six hours around Texas A&M University, and improved the short term prediction of RH at 850 mb. Low-level temperatures in the ICs and BCs from the NAM were much improved by the RAOB. WRF_OBR did not show much improvement over WRF_OB, but improvements were seen in the spatial maps at low-levels. Even though MADIS observations do include the low-levels, they are not widespread. This is the reason why WRF_OBR does show some improvements.

4.3 Subjective Verification

Three cases will be examined subjectively in this section, two convective cases, and a winter case. The convective cases were chosen based on how the simulations performed objectively with respect to precipitation for the WRF_OB assimilation strategy, while the winter case was selected because winter precipitation occurred close to and in College Station. The first convective case performed poorly with DA, the second convective case performed better with DA, and the winter case performed poorly with DA. Observed and simulated reflectivity is compared for all cases.

The first case had isolated convection and a line of convection to the northwest when a balloon was launched. Figure 4.25 shows the 10 mm CSI verification for this case.

One hour into the simulation at 19Z, Figure 4.26 shows that all simulations are very different than observed reflectivity. WRF_OBR is one of the better simulations and notably has less convection than WRF_OB that did not have the RAOB assimilated.

Moving on to the next day, the WRF_C and WRF_R simulations are similar, while WRF_OB and WRF_OBR are very different from each other. WRF_OBR looks much like the radar imagery for these hours and from Figure 4.27, it is evident that it characterizes the convection better than the others. The WRF_C and WRF_R simulations have developed a bow echo, but it is close to the line of convection. As the simulation time wraps up, the WRF_OB simulation continues to poorly represent convection, and WRF_R most resembles the precipitation at 12Z (Figure 4.28).

From this time to 18Z, nothing notable happens as most convection is gone in the observations and the models. WRF_OB, however, still generates more convection than the other simulations. In summary, WRF_OBR and WRF_R represent convection fairly well in the model, as the RAOB helps the WRF_OBR simulation suppress spurious convection versus the WRF_OB simulation.

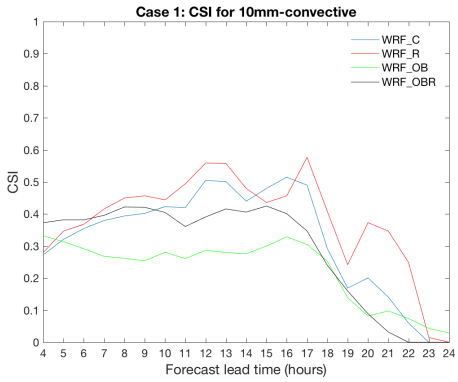

Figure 4.25: Verification of 10 mm rainfall for the simulation on 20 May 2011 at 18Z.

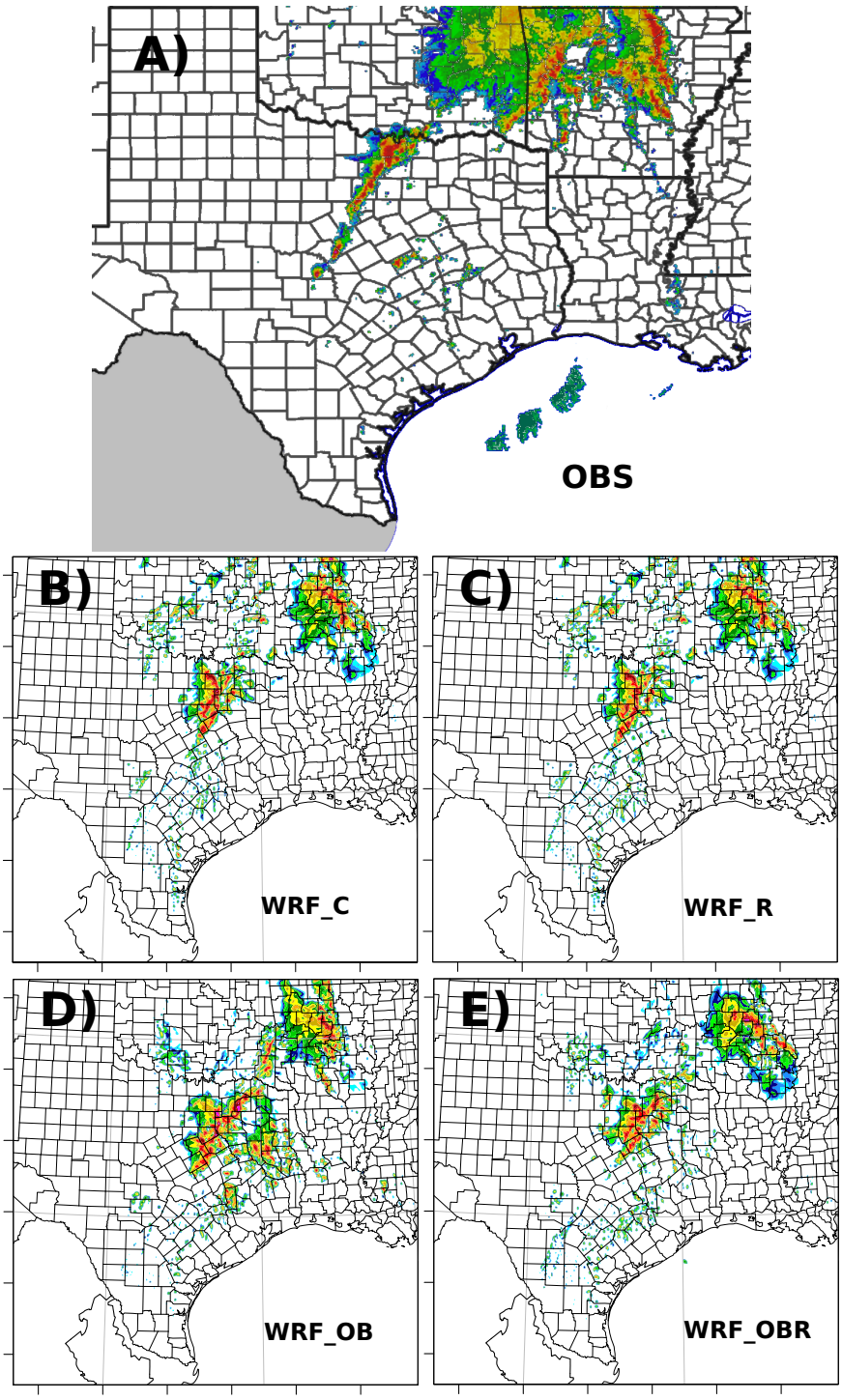

Figure 4.26: Observed and simulated reflectivity for 20 May, 2011 at 19Z. Panel A is the observed base reflectivity, while the other panels are simulated reflectivity for the simulations.

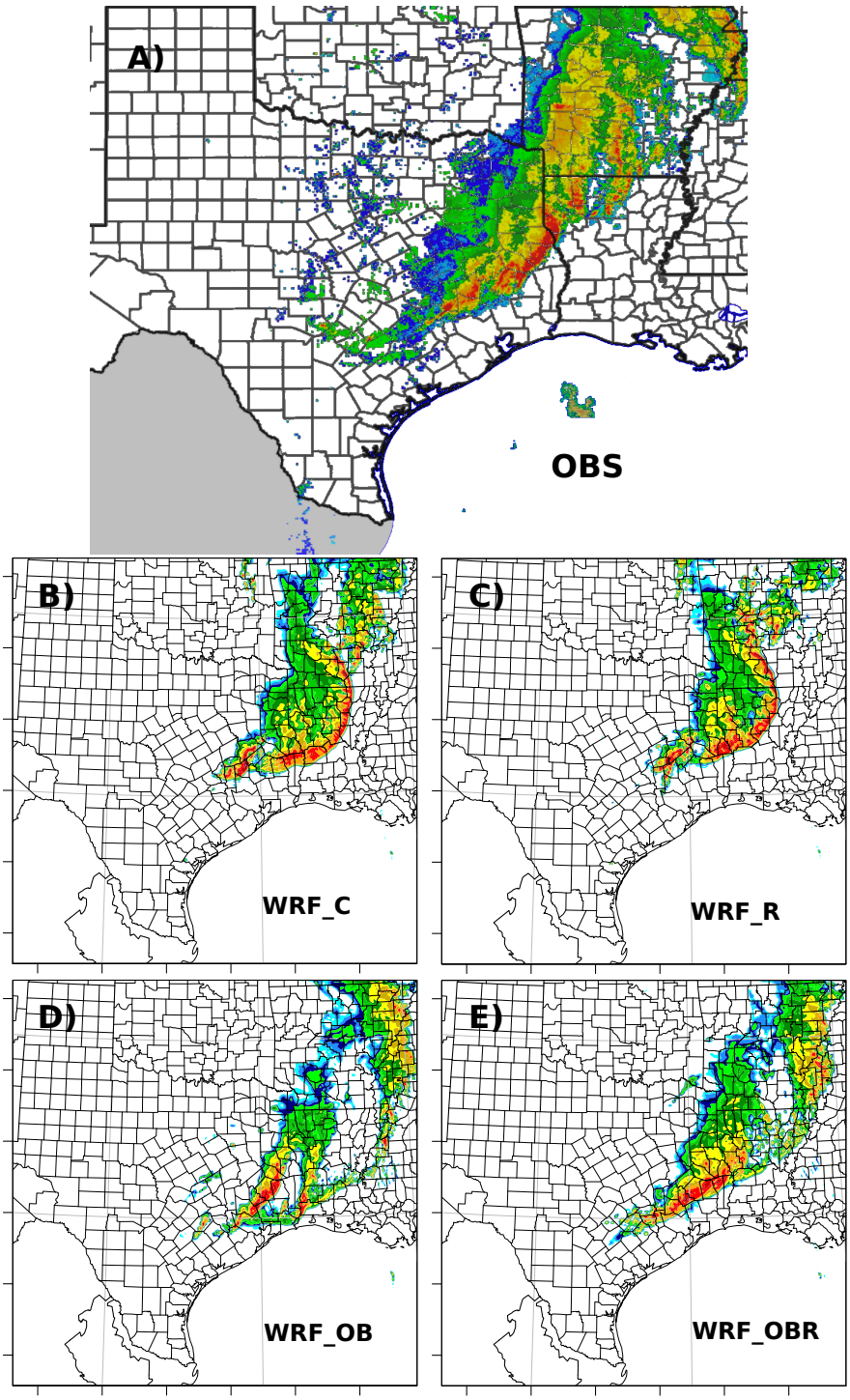

Figure 4.27: Observed and simulated reflectivity for 21 May, 2011 at 6Z. Panel A is the observed base reflectivity, while the other panels are simulated reflectivity for the simulations.

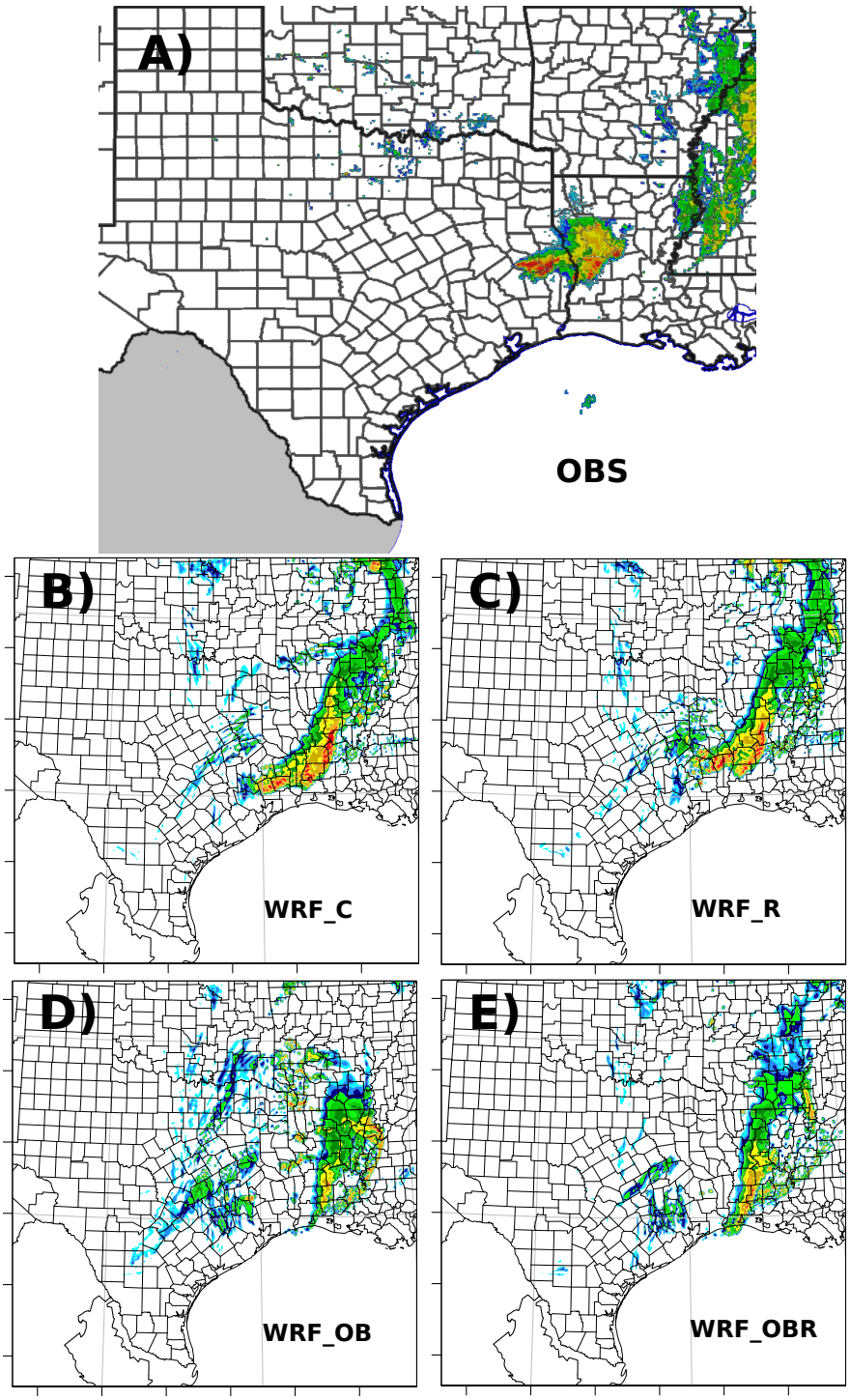

Figure 4.28: Observed and simulated reflectivity for 21 May, 2011 at 12Z. Panel A is the observed base reflectivity, while the other panels are simulated reflectivity for the simulations.

Another date was chosen for subjective verification, but this case performed relatively well when objective verification was done for WRF_OB. Figure 4.29 shows the objective verification at 10 mm.

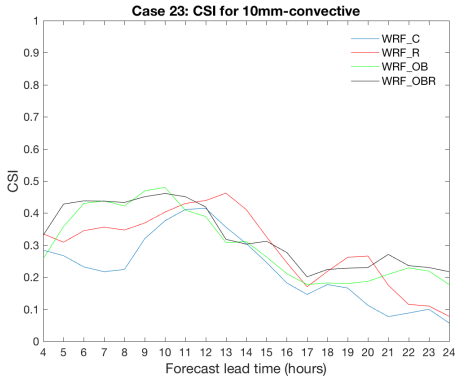

Figure 4.29: Verification of 10mm rainfall for the simulation on 29 April 2016 at 20Z.

Initially at forecast hour one, and similar to the previous case, no simulation represents convection well. WRF_OBR best represents the observed reflectivity for the first several hours, as the RAOB assimilation appears to have helped again. Figure 4.30 shows very different observations and simulations, but it is clear that WRF_OBR has captured the precipitation the best. WRF_R and WRF_OBR match the observations the best over the next several hours and this can be seen in Figure 4.31, 21 hours after simulation start time. WRF_OBR is again improved from WRF_OB because of the RAOB assimilation.

WRF_OBR performed the best out of all simulations for this case, capturing more features of the convection at later hours. It does have some spurious convection towards the end, but it positions convection better in LA. WRF_R also performed better than the WRF_C simulation, shifting precipitation closer to the observations.

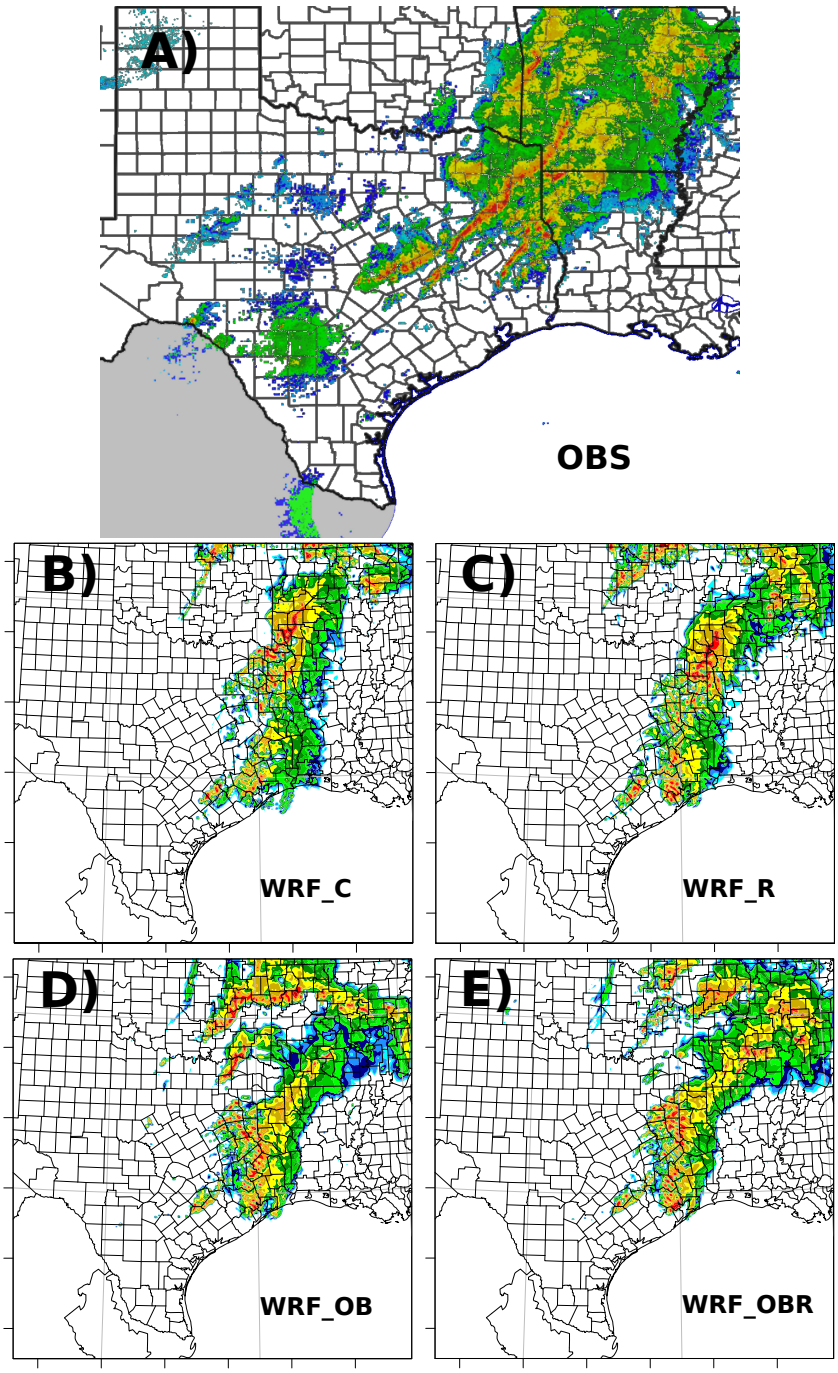

Figure 4.30: Observed and simulated reflectivity for 30 April, 2016 at 5Z. Panel A is the observed base reflectivity, while the other panels are simulated reflectivity for the simulations.

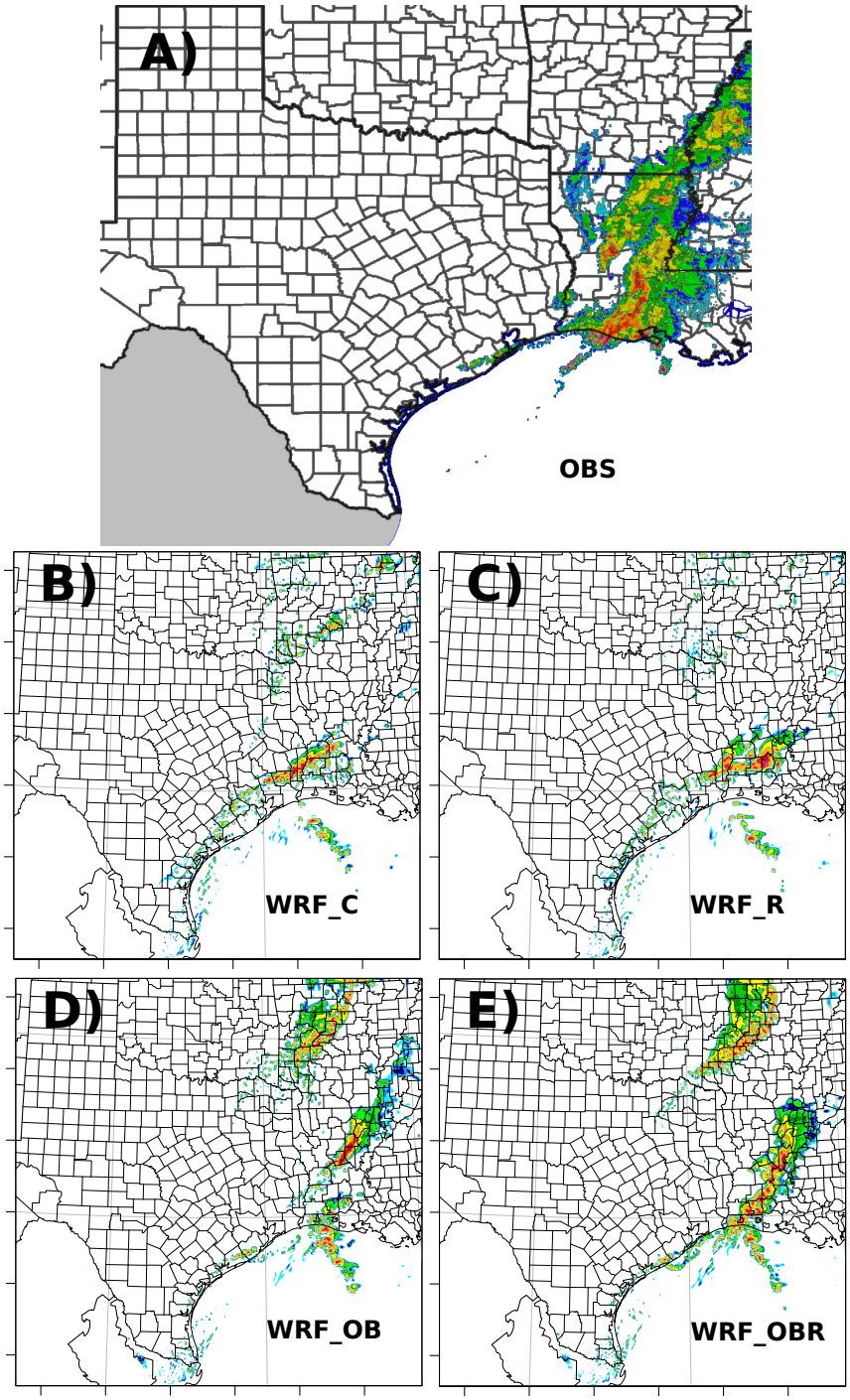

Figure 4.31: Observed and simulated reflectivity for 30 April, 2016 at 17Z. Panel A is the observed base reflectivity, while the other panels are simulated reflectivity for the simulations.

The last case is a winter one, where winter precipitation occurred across the state and into College Station. The analysis will also look at winter precipitation types during the simulation time. Figure 4.32 shows the poor objective performance of the case at 1 mm.

Simulations start out very different than observed precipitation. Initial position of the precipitation is poor for all simulations, but especially WRF_OB and WRF_OBR. Over the next several hours, no simulation captures the precipitation well, as they miss significant amounts of stratiform rain over eastern TX and LA (Figure 4.33). Freezing lines at the surface and 850 mb are overlaid on these plots in order to see how different a precipitation type forecast could be. Freezing at the surface for the WRF_OB and WRF_OBR simulations is much different than the two other simulations and makes a big difference in terms of winter precipitation type. Their results match up more closely with observations as well.

During the remaining 9 hours, poor representation of convection continues, and at 0Z the next day (Figure 4.34), all simulations basically converge. The temperature forecasts do not appear much different here. All simulations poorly characterized the rainfall, with the DA simulations performing poorly. Even though the rainfall was not represented well, the temperature forecasts were better overall with WRF_OB and WRF_OBR that assimilated MADIS data.

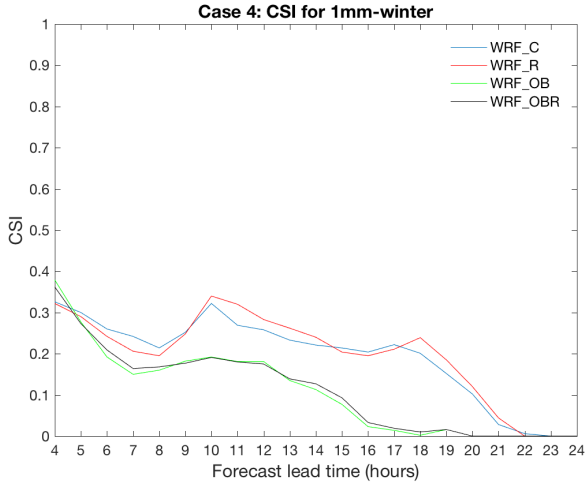

Figure 4.32: Verification of 1mm rainfall for the simulation on 24 January 2014 at 0Z.

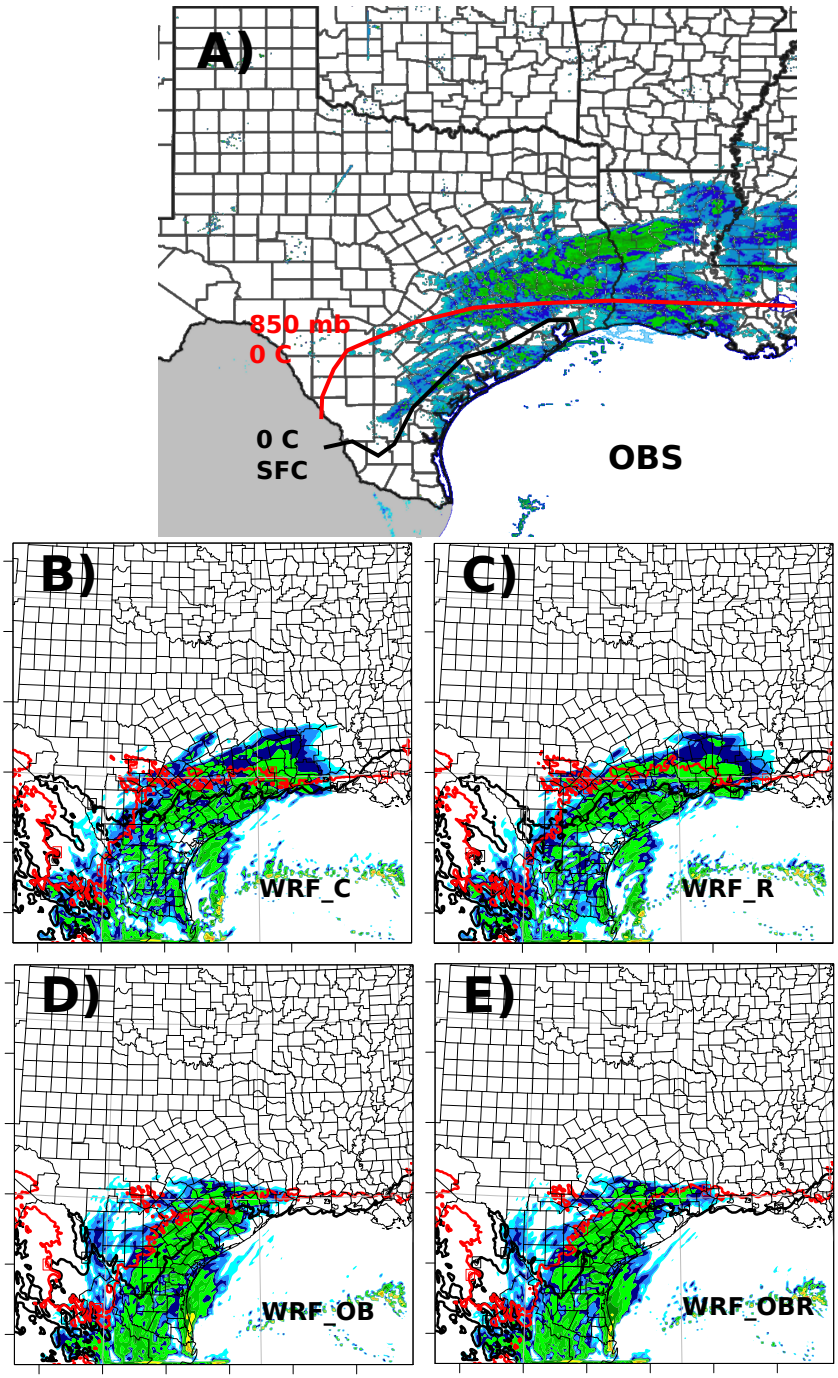

Figure 4.33: Observed and simulated reflectivity for 24 January, 2014 at 9Z. Panel A is the observed base reflectivity, while the other panels are simulated reflectivity for the simulations. Panel A has the 0 C line for 9Z and the 850 mb 0 degree C line at 12Z. For the simulation plots, the black solid line is the 0 degree C line at the surface, while the red line is the 0 degree C line at approximately 850 mb. The 0 C line for the surface was a hand analysis of surface observations at 9Z, while the 12Z 850 mb 0 C line was drawn from the SPC Upper-Air Archive

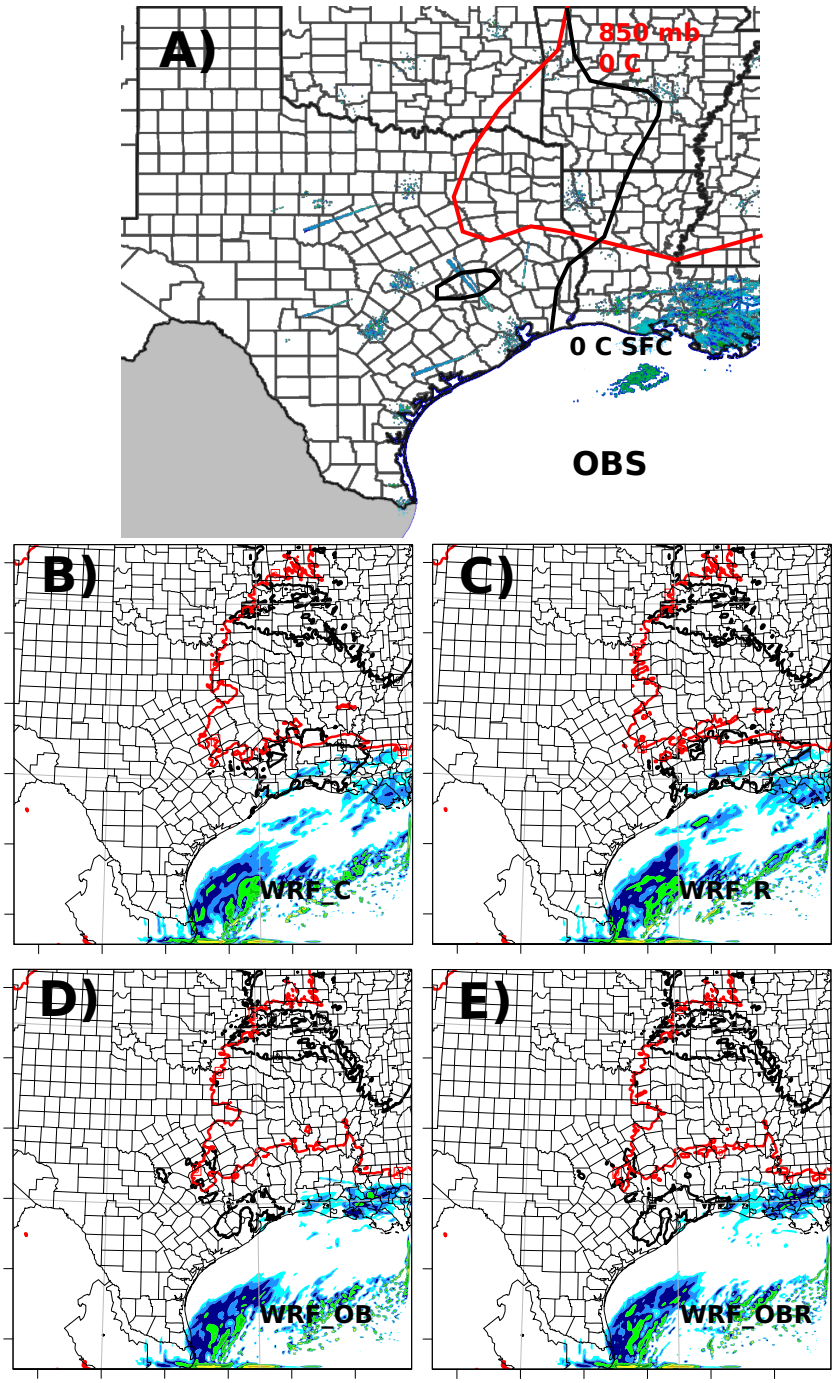

Figure 4.34: Observed and simulated reflectivity for 25 January, 2014 at 0Z. Panel A has the 0 degreee C line for 0Z and the 850 mb 0 degree C line at 0Z as well. For the simulation plots, the black solid line is the 0 degree C line at the surface, while the red line is the 0 degree C line at approximately 850 mb. The 0 C line for the surface was a hand analysis of surface observations at 0Z, while the 0Z 850 mb 0 C line was drawn from the SPC Upper-Air Archive

4.3.1 Subjective Verification Summary

For the two convective cases, WRF_R and WRF_OBR did perform slightly better than the other simulations. In the first case, WRF_C had a better objective precipitation score than WRF_OBR for much of the time, but the structure of the precipitation was better for WRF_OBR for parts of the simulation as well. The second case was a little different since they all performed poorly, but WRF_OBR was superior of the four simulations for the majority of the time. The winter case had poor performance in all simulations in terms of precipitation, as the WRF_C simulation outperformed WRF_OB and WRF_OBR, and was very similar to WRF_R. As far as temperature is concerned for the winter case, WRF_OB and WRF_OBR performed well, having little difference between the two simulations.

4.4 Limited DA Testing

A different DA strategy was employed for two cases. Since a large amount and variety of data types were assimilated for the WRF_OB and WRF_OBR simulations, some of the data could be of poor quality. Introduction of the worst quality data will be eliminated in WRF-DA, but some inferior observations could pass quality control, making their way into the assimilation. For two dates, May 20th, 2011 and May 5th, 2015, limited DA was done to test the simulation performance with fewer observations. Meso-net observations, satellite observed winds, and profilers such as MAP and NPN were excluded from DA. By excluding these data types, which are generally of lower quality than the others, we can see some of the forecast sensitivity to these observations. The simulation results were briefly examined using an objective verification of precipitation.

The first performed poorly with the original DA strategy, but performed much better when assimilating limited data. Figure 4.35 shows the performance of the

limited observation DA for a precipitation threshold of 10 mm. All three variables improve at all forecast hours. Poor quality observations in the MADIS dataset, regardless of the quality control in WRF-DA, could be introduced into the assimilation. The second case was less sensitive to the limited DA, but it did have higher scores in general. One other possible reason for the results is the R matrix for observations in WRF-DA. This matrix assumes that each data type has the same error. Mesonet observations, however, for example, are from a variety of providers, so they likely do not have the same errors among stations. In DA, no observation should be unused, but unfortunately, even if you know the specific error of a station, WRF-DA does not yet have capabilities to use this info. The RAOB also has slightly more weight with this strategy because less data is assimilated in this case, potentially having a more positive affect on the forecast.

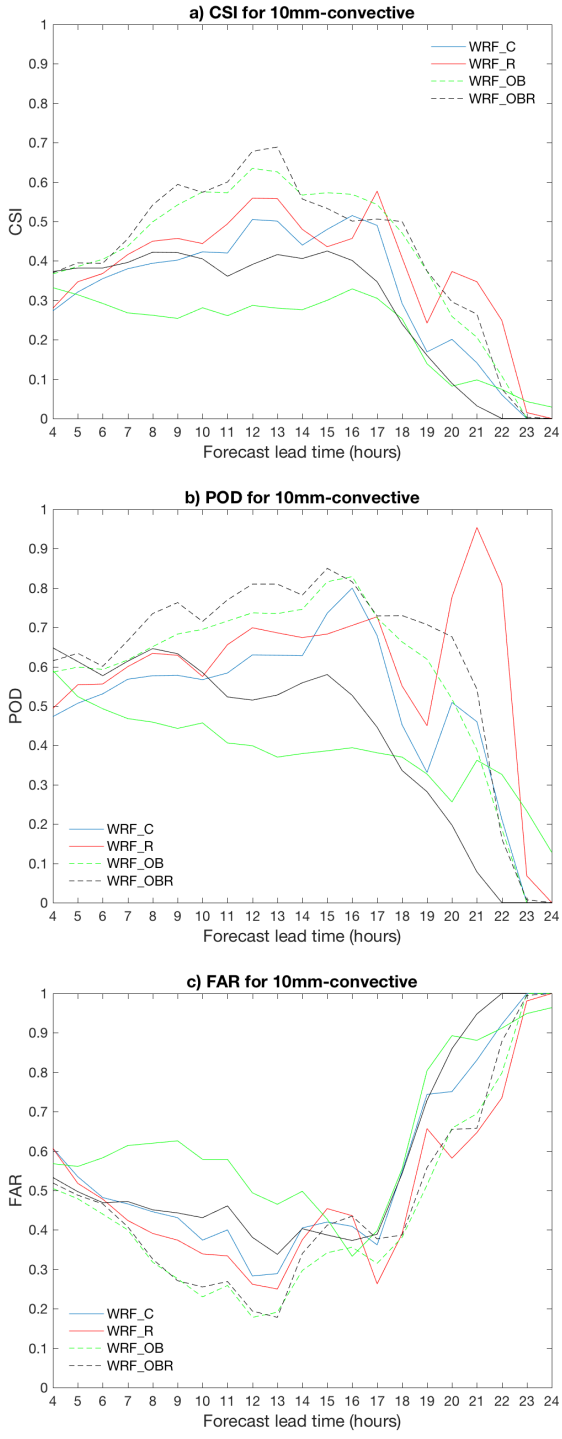

Figure 4.35: CSI, POD, and FAR for the May 20th, 2011 and the 10 mm precipitation threshold. The dashed color lines represent the limited DA simulations and the solid lines are for the standard cases.

5. SUMMARY AND CONCLUSIONS

At the beginning of this paper, it was hypothesized that a RAOB in a data sparse region has potential benefit to NWP in a high-resolution convection allowing model. These RAOBs were useful to forecasters, who use a variety of upper air data for prediction of weather. Upper air data from RAOBs has long been used in weather models and is known to provide benefit to forecasts, even when other data is available (Benjamin et al., 2010). Research in this study determined the NWP benefit of on-demand RAOBs using a high-resolution WRF model forecast. Objective and subjective spatial impact was evaluated for College Station and areas downstream.

DA included the single RAOBs as well as standard observations. One reason MADIS observations were included in the assimilation process was to emulate a real-time DA. The other reason was to test the impact of assimilating a single RAOB when a significant amount of data was already assimilated. WRF_R and WRF_OBR cases were only different from WRF_C and WRF_OB were only different in that they assimilated a single on-demand RAOB from Texas A&M. Differences in the simulation from a single RAOB assimilation was generally highest for WRF_R, and lowest for WRF_OBR. This was expected since WRF_OB assimilated MADIS observations while WRF_C only started with ICs and BCs from the NAM.

Verification of WRF_R showed some improvements in precipitation, around and downstream of College Station at late forecast hours in simulations. This was pronounced for rainfall thresholds of 25 mm over a three hour period. Characteristics of rainfall also improved in the two cases that were examined subjectively. When comparing atmospheric variables for the convective cases, there was not a noticeable advantage except in winter cases. For short forecasts in particular, the prediction

of surface and lower-level temperature was significantly better than the WRF_C simulation. This is important as temperature variations of 0.5 C or 1 C can dictate whether frozen or liquid precipitation falls. In addition, benefit was concentrated around College Station in the spatial maps.

WRF_OBR and WRF_OB verifications were both very similar in verification for precipitation and other atmospheric variables. WRF_OB assimilates a significant amount of data such as AirReps, profiler data, surface observations, and satellite observations of winds. When the RAOB was assimilated for WRF_OBR, the many assimilated observations had already improved the ICs, causing the RAOB to have a smaller, often neglibible influence on the forecast. Assimilation of MADIS data in the WRF_OB case improved convective precipitation forecasts compared to the control case WRF_C, and led to somewhat lower forecast errors. Generally, simulated reflectivity from the subjective verification of convective precipitation showed WRF_OBR to outperform WRF_OB, as it revealed a more representative structure of the precipitation. Assimilation of MADIS data improved winter surface temperature forecasts significantly, but was detrimental to the precipitation forecast for essentially every winter case. Further assimilation of the RAOB in the WRF_OBR case improved lower-level temperature forecasts modestly in the winter.

RAOB data from Texas A&M showed modest NWP forecast impact, particularly for precipitation characterization and intensity for convective cases and low-level temperatures, particularly in the winter. Impact and benefit was limited for the whole WRF model domain, while College Station and regions nearby received the most positive benefit. Not only does NWP show some benefit with these RAOBs, but this has been the experience of forecasters at local NWS offices and the SPC. Conlee et al. (2014) noted that local NWS offices received the benefit in forecasting convection, but most benefit came from the RAOBs in winter cases, improving their

understanding of low-level temperatures for precipitation type forecasting. RAOB launches from Texas A&M continue today; these findings encourage future support.

Future work could include an implementation of a real-time data assimilation system at Texas A&M that could incorporate the RAOB and MADIS data. Some work should investigate sensitivity of WRF-DA to different data types, especially for winter cases. From the small amount of testing done, there is a large amount of sensitivity to certain observations besides RAOBs. The MADIS data are known to have data types of varying quality, which if characterized properly, real-time DA could be implemented.

REFERENCES

Barker, D., Huang, X.-Y., Liu, Z., Auligné, T., Zhang, X., Rugg, S., Ajjaji, R., Bourgeois, A., Bray, J., and Chen, Y. (2012). The weather research and forecasting model's community variational/ensemble data assimilation system:WRFDA. *Bulletin of the American Meteorological Society*, 93(6):831-843.

Barnes, S. L. (1964). A technique for maximizing details in numerical weather map analysis. *Journal of Applied Meteorology*, 3(4):396-409.

Benjamin, S. G., Dvnyi, D., Weygandt, S. S., Brundage, K. J., Brown, J. M., Grell, G. A., Kim, D., Schwartz, B. E., Smirnova, T. G., and Smith, T. L. (2004). An hourly assimilation-forecast cycle: The RUC. *Monthly Weather Review*, 132(2):495-518.

Benjamin, S. G., Jamison, B. D., Moninger, W. R., Sahm, S. R., Schwartz, B. E., and Schlatter, T. W. (2010). Relative short-range forecast impact from aircraft, profiler, radiosonde, vad, GPS-PW, metar, and mesonet observations via the RUC hourly assimilation cycle. *Monthly Weather Review*, 138(4):1319-1343.

Bergthórsson, P. and Döös, B. R. (1955). Numerical weather map analysis. *Tellus*, 7(3):329-340.

Bouttier, F. and Kelly, G. (2001). Observing-system experiments in the ECMWF 4DVar data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 127(574):1469-1488.

Brown, B. G., Gotway, J. H., Bullock, R., Gilleland, E., Fowler, T., Ahijevych, D., and Jensen, T. (2009). The model evaluation tools (MET): Community tools

for forecast evaluation. In *Preprints, 25th Conf. on International Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology, Phoenix, AZ, Amer. Meteor. Soc. A*, volume 9, page 6.

Buehner, M. (2005). Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting. *Quarterly Journal of the Royal Meteorological Society*, 131(607):1013-1043.

Charney, J. G., Fjörtoft, R., and von Neumann, J. (1950). Numerical integration of the barotropic vorticity equation. *Tellus*, 2(4):237-254.

Chen, F. and Dudhia, J. (2001). Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. part II: Preliminary model validation. *Monthly Weather Review*, 129(4):587-604.

Coniglio, M. C., Jr, J. C., Marsh, P. T., and Kong, F. (2013). Verification of convection-allowing WRF model forecasts of the planetary boundary layer using sounding observations. *Weather and Forecasting*, 28(3):842-862.

Conlee, D., Wood, L., and Sodowsky, R. (2014). On-demand radiosonde observations to aid severe and hazardous weather forecasting in the southeast Texas upper-air sparse region. In *18th Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface*, Austin, TX.

Cressman, G. P. (1959). An operational objective analysis system. *Monthly Weather Review*, 87(10):367-374.

Deb, S., Kishtawal, C., Pal, P., and Joshi, P. (2008). A modified tracer selection and tracking procedure to derive winds using water vapor imagers. *Journal of Applied Meteorology and Climatology*, 47(12):3252-3263.

Divakarla, M. G., Barnet, C. D., Goldberg, M. D., McMillin, L. M., Maddy, E., Wolf, W., Zhou, L., and Liu, X. (2006). Validation of atmospheric infrared sounder temperature and water vapor retrievals with matched radiosonde measurements and forecasts. *Journal of Geophysical Research: Atmospheres*, 111(D9).

Done, J., Davis, C. A., and Weisman, M. (2004). The next generation of NWP: Explicit forecasts of convection using the weather research and forecasting (WRF) model. *Atmospheric Science Letters*, 5(6):110-117.

Dong, J., Xue, M., and Droegemeier, K. (2011). The analysis and impact of simulated high-resolution surface observations in addition to radar data for convective storms with an ensemble kalman filter. *Meteorology and Atmospheric Physics*, 112(1-2):41-61.

Ek, M., Mitchell, K., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., and Tarpley, J. (2003). Implementation of noah land surface model advances in the national centers for environmental prediction operational mesoscale eta model. *Journal of Geophysical Research: Atmospheres*, 108(D22).

Evensen, G. (2003). The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, 53(4):343-367.

Evensen, G. (2009). *Data assimilation: The ensemble Kalman filter*. Springer Science & Business Media.

Fisher, M. (2003a). Background error covariance modelling. In *Seminar on Recent Development in Data Assimilation for Atmosphere and Ocean*, pages 45-63.

Fisher, M. (2003b). *Estimation of entropy reduction and degrees of freedom for*

signal for large variational analysis systems. European Centre for Medium-Range Weather Forecasts, Reading, UK.

Gelaro, R., Langland, R. H., Pellerin, S., and Todling, R. (2010). The thorpex observation impact intercomparison experiment. *Monthly Weather Review*, 138(11):4009-4025.

Gilchrist, B. and Cressman, G. P. (1954). An experiment in objective analysis. *Tellus*, 6(4):309-318.

Goldberg, M. D., Qu, Y., McMillin, L. M., Wolf, W., Zhou, L., and Divakarla, M. (2003). AIRS near-real-time products and algorithms in support of operational numerical weather prediction. *Geoscience and Remote Sensing, IEEE Transactions on*, 41(2):379-389.

Gutman, S. I., Sahm, S. R., Benjamin, S. G., Schwartz, B. E., Holub, K. L., Stewart, J. Q., and Smith, T. L. (2004). Rapid retrieval and assimilation of ground based GPS precipitable water observations at the NOAA forecast systems laboratory: Impact on weather forecasts. *Journal of the Meteorological Society of Japan, Series 2*, 82(1B):351-360.

Ha, J.-H., Kim, H.-W., and Lee, D.-K. (2011). Observation and numerical simulations with radar and surface data assimilation for heavy rainfall over central Korea. *Advances in Atmospheric Sciences*, 28(3):573-590.

Hoke, J. E. and Anthes, R. A. (1976). The initialization of numerical models by a dynamic-initialization technique. *Monthly Weather Review*, 104(12):1551-1556.

Hollingsworth, A. and Lönnberg, P. (1986). The statistical structure of shortrange

forecast errors as determined from radiosonde data. part I: The wind field. *Tellus A*, 38(2):111--136.

Hong, S.-Y., Noh, Y., and Dudhia, J. (2006). A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, 134(9):2318-2341.

Hou, T., Kong, F., Chen, X., and Lei, H. (2013). Impact of 3DVAR data assimilation on the prediction of heavy rainfall over southern China. *Advances in Meteorology*, 2013.

Hu, X.-M., Nielsen-Gammon, J., and Zhang, F. (2010). Evaluation of three planetary boundary layer schemes in the WRF model. *Journal of Applied Meteorology & Climatology*, 49(9):1831-1844.

Huang, X.-Y. and Vedel, H. (2003). An introduction to data assimilation. In *SAF Training Workshop: 2nd GRAS SAF User Workshop, EUMETSAT, Darmstadt, Germany*, pages 25-37.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D. (2008). Radiative forcing by longlived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research: Atmospheres*, 113(D13).

Inai, Y., Hasebe, F., Shimizu, K., and Fujiwara, M. (2009). Correction of radiosonde pressure and temperature measurements using simultaneous GPS height data. *SOLA*, 5:109-112.

Inai, Y., Shiotani, M., Fujiwara, M., Hasebe, F., and Vmel, H. (2015). Altitude mises-

timation caused by the vaisala rs80 pressure bias and its impact on meteorological profiles. *Atmospheric Measurement Techniques Discussions*, 8(2):2191-2222.

Johnson, A., Wang, X., Kong, F., and Xue, M. (2013). Object-based evaluation of the impact of horizontal grid spacing on convection-allowing forecasts. *Monthly Weather Review*, 141(10):3413-3425.

Kain, J. S., Xue, M., Coniglio, M. C., Weiss, S. J., Kong, F., Jensen, T. L., Brown, B. G., Gao, J., Brewster, K., and Thomas, K. W. (2010). Assessing advances in the assimilation of radar data and other mesoscale observations within a collaborative forecasting-research environment. *Weather and Forecasting*, 25(5):1510-1521.

Kalnay, E. (2003). *Atmospheric modeling, data assimilation and predictability*. Cambridge University Press, Cambridge, England.

Kelly, G., McNally, A., Thpaut, J., and Szyndel, M. (2004). Observing system experiments of all main data types in the ECMWF operational system. In *3rd WMO Numerical Weather Prediction OSE Workshop, Alpbach, Austria, WMO, Technical Report*, volume 1228, pages 32-36.

Kelly, G., Thpaut, J., Buizza, R., and Cardinali, C. (2007). The value of observations. I: Data denial experiments for the Atlantic and the Pacific. *Quarterly Journal of the Royal Meteorological Society*, 133(628):1803-1815.

Kistler, R. E. (1974). *A study of data assimilation techniques in an autobarotropic, primitive equation, channel model*. PhD dissertation, Penn State University, State College, PA.

Langland, R. H. (2005). Issues in targeted observing. *Quarterly Journal of the Royal Meteorological Society*, 131(613):3409-3425.

Lönnberg, P. and Hollingsworth, A. (1986). The statistical structure of shortrange forecast errors as determined from radiosonde data part II: The covariance of height and wind errors. *Tellus A*, 38(2):137-161.

Lynch, P. (2006). *The emergence of numerical weather prediction: Richardson's dream*. Cambridge University Press, Cambridge, England.

Madsen, H. and Canizares, R. (1999). Comparison of extended and ensemble Kalman filters for data assimilation in coastal area modelling. *International Journal for Numerical Methods in Fluids*, 31(6):961-981.

Mass, C. F., Ovens, D., Westrick, K., and Colle, B. A. (2002). Does increasing horizontal resolution produce more skillful forecasts? *Bulletin of the American Meteorological Society*, 83(3).

Mellor, G. L. and Yamada, T. (1982). Development of a turbulence closure model for geophysical fluid problems. *Reviews of Geophysics*, 20(4):851-875.

Michalakes, J., Chen, S., Dudhia, J., Hart, L., Klemp, J., Middlecoff, J., and Skamarock, W. (2001). Development of a next generation regional weather research and forecast model. In *Developments in Teracomputing: Proceedings of the Ninth ECMWF Workshop on the use of high performance computing in meteorology*, volume 1, pages 269-276. World Scientific Conference.

Miller, P. A., Barth, M., Benjamin, L., Artz, R., and Pendergrass, W. (2005a). The meteorological assimilation and data ingest system (MADIS): Providing value-added observations to the meteorological community. In *21st Conference on Weather Analysis and Forecasting*.

Miller, P. A., Barth, M. F., and Benjamin, L. A. (2005b). An update on MADIS

observation ingest, integration, quality control, and distribution capabilities. In *Preprints, 21st Int. Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology, San Diego, CA, Amer. Meteor. Soc., CD-ROM J.*

Nakanishi, M. and Niino, H. (2004). An improved melloryamada level-3 model with condensation physics: Its design and verification. *Boundary-Layer Meteorology*, 112(1):1-31.

Nakanishi, M. and Niino, H. (2009). Development of an improved turbulence closure model for the atmospheric boundary layer. *Journal of the Meteorological Society of Japan, Series 2*, 87(5):895-912.

Navon, I. M. (2009). *Data assimilation for numerical weather prediction: A review*, pages 21-65. Data assimilation for atmospheric, oceanic and hydrologic applications. Springer.

Otkin, J., Huang, H.-L., and Seifert, A. (2006). A comparison of microphysical schemes in the WRF model during a severe weather event. In *Papers delivered at 7th WRF Users Workshop, Boulder, CO, USA*, pages 19-22.

Panofsky, R. (1949). Objective weather-map analysis. *Journal of Meteorology*, 6(6):386-392.

Parrish, D. F. and Derber, J. C. (1992). The national meteorological center's spectral statistical-interpolation analysis system. *Monthly Weather Review*, 120(8):1747-1763.

Pleim, J. E. (2007a). A combined local and nonlocal closure model for the atmo-

spheric boundary layer. part I: Model description and testing. *Journal of Applied Meteorology and Climatology*, 46(9):1383-1395.

Pleim, J. E. (2007b). A combined local and nonlocal closure model for the atmospheric boundary layer. part II: Application and evaluation in a mesoscale meteorological model. *Journal of Applied Meteorology and Climatology*, 46(9):1396-1409.

Pondeca, M. S. D., Manikin, G. S., DiMego, G., Benjamin, S. G., Parrish, D. F., Purser, R. J., Wu, W.-S., Horel, J. D., Myrick, D. T., and Lin, Y. (2011). The real-time mesoscale analysis at NOAA's national centers for environmental prediction: Current status and development. *Weather and Forecasting*, 26(5):593-612.

Rennie, S. J., Dance, S. L., Illingworth, A. J., Ballard, S., and Simonin, D. (2011). 3D-Var assimilation of insect-derived doppler radar radial winds in convective cases using a high-resolution model. *Monthly Weather Review*, 139(4):1148-1163.

Richardson, L. (1922). *Weather prediction by numerical methods*. Cambridge University Press, Cambridge, England.

Skamarock, W. (2008). A description of the advanced research WRF, version 3. *National Center for Atmospheric Research Technical Note*.

Stauffer, R., Morris, G., Thompson, A., Joseph, E., and Coetzee, G. (2013). Propagation of radiosonde pressure sensor errors to ozonesonde measurements. *Atmospheric Measurement Techniques Discussions*, 6:7771-7810.

Sun, J., Trier, S. B., Xiao, Q., Weisman, M. L., Wang, H., Ying, Z., Xu, M., and Zhang, Y. (2012). Sensitivity of 0-12-h warm-season precipitation forecasts over the central United States to model initialization. *Weather and Forecasting*, 27(4):832-855.

Sun, J. and Wang, H. (2013). Radar data assimilation with WRF 4D-Var. part ii: Comparison with 3D-Var for a squall line over the US Great Plains. *Monthly Weather Review*, 141(7):2245-2264.

Thompson, G., Field, P. R., Rasmussen, R. M., and Hall, W. D. (2008). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. part II: Implementation of a new snow parameterization. *Monthly Weather Review*, 136(12):5095-5115.

Whitaker, J. S., Compo, G. P., Wei, X., and Hamill, T. M. (2004). Reanalysis without radiosondes using ensemble data assimilation. *Monthly Weather Review*, 132(5):1190-1200.

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences*, volume 100. Academic Press, Cambridge, MA.

Yesubabu, V., Srinivas, C. V., Langodan, S., and Hoteit, I. (2015). Predicting extreme rainfall events over Jeddah, Saudi Arabia: Impact of data assimilation with conventional and satellite observations. *Quarterly Journal of the Royal Meteorological Society*.

Zavodsky, B., Case, J. L., Gotway, J. H., White, K., Medlin, J., Wood, L., and Radell, D. (2014). Development and implementation of dynamic scripts to support local model verification at national weather service weather forecast offices. *NASA*.